

DOCUMENT RESUME

ED 075 715

AC 014 354

AUTHOR Steele, Sara M.
TITLE Contemporary Approaches to Program Evaluation and Their Implications for Evaluating Programs for Disadvantaged Adults.
INSTITUTION Syracuse Univ., N.Y. ERIC Clearinghouse on Adult Education.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE 19 May 73
NOTE 266p.
EDRS PRICE MF-\$0.65 HC-\$9.87
DESCRIPTORS *Adult Education Programs; Bibliographies; *Disadvantaged Groups; *Evaluation Techniques; Guides; *Program Evaluation; Taxonomy

ABSTRACT

Designed as a reference to contemporary evaluation approaches, this monograph brings together a variety of old and new frameworks and ideas about evaluation and shows how they are related to one another. Section I provides an overview of the contemporary evaluation scene. Section III presents summaries of over 50 approaches to evaluation from a variety of social fields, some preliminary sketches and others extensively developed. The middle section provides guidelines on how to sort these out--how to choose and use evaluation approaches. Appendixes give condensed contemporary program evaluation approaches, evaluation approaches categorized by field of origin, and a try at developing a taxonomy of program evaluation. There is an extensive bibliography, an index of authors, and an index of approaches. (Author)

ED 075715

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE

OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

CONTEMPORARY APPROACHES TO PROGRAM EVALUATION

and

THEIR IMPLICATIONS FOR EVALUATING
PROGRAMS FOR DISADVANTAGED ADULTS

SARA M. STEELE

ERIC CLEARINGHOUSE ON ADULT EDUCATION
107 Roney Lane, Syracuse, N.Y. 13210

AC014354

This publication was prepared pursuant to a contract with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their judgments in professional and technical matters. Points of view or opinions do not, therefore, necessarily represent official Office of Education position or policy.

ABSTRACT

Designed as a reference to contemporary evaluation approaches, this monograph brings together a variety of old and new frameworks and ideas about evaluation and shows how they are related to one another. Section I provides an overview of the contemporary evaluation scene. Section III presents summaries of over 50 approaches to evaluation from a variety of social fields, some preliminary sketches and others extensively developed. The middle section provides guidelines on how to sort these out--how to choose and use evaluation approaches. Appendixes give condensed contemporary program evaluation approaches, evaluation approaches categorized by field of origin, and a try at developing a taxonomy of program evaluation. There is an extensive bibliography, an index of authors, and an index of approaches.

FOREWORD

Evaluation has come a long way in education during the past five years. Suddenly, it seems, everyone responsible for programs has become conscious of evaluation within the context of accountability, as well as improvement of their programs. In addition, research and development projects sponsored by the Federal Government, and other agencies, have had provisions for evaluations included in the contracts.

At the same time, increasing numbers of scholars have been re-examining the existing evaluation models, have been modifying them and proposing new versions. To the credit of the field, adult educators have been among those in the forefront adding to the development of evaluation models.

Perhaps the most pressing need for evaluation models in adult education has been in the area of programs for disadvantaged adults. Recognizing this need, the ERIC Clearinghouse on Adult Education took the initiative to bring together the many contemporary approaches to program evaluation as they can be applied to programs for disadvantaged adults. To this end, the Clearinghouse commissioned Dr. Sara M. Steele to produce a monograph on this topic.

After a year's concentrated work, Dr. Steele and her team of colleagues at the University of Wisconsin, produced this monograph. Nearly fifty scholars and practitioners provided critical comments on the various drafts of this monograph. The end product is a rather comprehensive compilation of evaluation models suitable for programs

for disadvantaged adults, plus some practical guidelines for applying the models to various situation which may confront a practitioner.

ERIC/AE takes pride in presenting this document, and is grateful to Sara Steele for her contribution to the field.

May 19, 1973

Stanely M. Grabowski
Director
ERIC Clearinghouse on Adult Education

PREFACE

This monograph is written for those responsible for designing evaluation strategy for an institution or agency. In some cases, this is the person who conducts the program. In most instances, however, it's an administrator or his assistants. Most of the intended audience has some knowledge of much of the material included. This publication, then, is a source of review and a reference for them. It may also be of use to other program personnel.

Although it was contracted for in terms of evaluating recent social programs, much of the content originated in other fields and most of it is applicable regardless of the type of program that is being evaluated.

The monograph is a major team effort. The help of the following people is acknowledged and greatly appreciated:

- the Wisconsin production team: Merrill Ewert, project assistant; Colleen Schuh, editor; Alice Durnford, secretary; Linda Weeden, final typist; Barbara Schutz, Charlene Durairaj, and Ruth Jacobson, typists; Beverly Cunningham, administrative secretary.
- Dr. Patrick Boyle and other administrators and colleagues here at the University of Wisconsin who arranged changes in responsibilities so that the contract could be undertaken and who gave support and encouragement throughout the project.
- the authors who created and refined the concepts and ideas summarized in the monograph and the publishers who brought those ideas into circulation.
- program personnel from throughout the United States and Canada who fitted reviewing the first draft into their extremely busy schedules.
- professors of adult and extension education who gave impetus to the development of the classification structure.
- the Office of Education and ERIC Clearinghouse on Adult Education which contracted with the University of Wisconsin for the manuscript and, in particular, Dr. Stanley Grabowski, director of ERIC, who provided the opportunity for proposing the idea for the manuscript and then guided its development.

CONTENTS

<i>Section</i>	<i>Page</i>
Introduction	1
I Looking at the Contemporary Evaluation Scene	3
Needs and Issues in Evaluating Programs for Disadvantaged Adults	3
Variance in Viewpoints and Needs	4
Multiple Units of Programs	9
Influence of State and Federal Funding	10
Inherentness of Conflict	12
Lack of Past Operating History	13
Relationship to Established Programs	14
Assumptions About Successful Programs	15
Contemporary Ideas About Program Evaluation	20
Program Evaluation Is a Process Rather Than a Procedure. It Is Generic Rather Than Specific	22
Program Evaluation Is More Than Examining the Attainment of Objectives	23
Program Evaluation Is More Than Evaluating the Results of a Program	26
Program Evaluation Is More Than Instructional Evaluation	28
Program Evaluation Is Different from Evaluative Research and Program Research	30
Program Evaluation Is a Management Tool	35
Program Evaluation Is People Centered	37
Current Approaches to Evaluation	39
The Work That Lies Ahead	48

II	Matching Evaluation Approaches to Needs	52
	Using Evaluation Approaches	52
	Expanding Your Ideas About Programming and Program Evaluation	52
	Locating Sources	53
	Providing You with Approaches You Can Adapt to Your Own Situation	53
	Providing You with Support and Ideas for Building Your Own Approaches	55
	Choosing Evaluation Approaches	55
	Types of Evaluation Frameworks	56
	Use of an Approach to Deal with a Specific Problem	59
III	Summaries of Contemporary Program Evaluation Approaches	75
	Group 1: Evaluation As Input Into Decision Making	76
	Group 2: Evaluation of Program Parts	101
	Group 3: Evaluation--Kinds of Data; Types of Activities	119
	Group 4: Evaluation Processes	134
	Group 5: Results--Attainment of Objectives	152
	Group 6: Results--Evaluation of Outcomes and Effects	190
	APPENDIX A Condensed Contemporary Program Evaluation Approaches	220
	APPENDIX B Evaluation Approaches Categorized by Field of Origin	234
	APPENDIX C A Try at Developing a Taxonomy of Program Evaluation Approaches	236
	BIBLIOGRAPHY	242
	INDEX OF AUTHORS	251
	INDEX OF APPROACHES	254

LIST OF TABLES

<i>Table</i>		<i>Page</i>
1	Contemporary Definitions of Evaluation	21
2	Grouping of Approaches to Program Evaluation	50
3	Commonly Needed Evaluation Frameworks	57
4	Matching Evaluation Approaches to Problems and Needs	61
C-1	A System for Classifying Evaluation Approaches	238
C-2	Location Within the Classification System of Adult and Extension Education and Other Major Approaches	240

INTRODUCTION

In the last 10 years many programs have emerged, such as adult basic education, manpower training, applied nutrition, human resource development, citizen education and leadership development, rehabilitation, that need to handle evaluation effectively and efficiently. The last five years have produced a host of ideas about what program evaluation is, what it should do, and how it should be done. Ideas about evaluation are in flux and transition. Yet, the average person responsible for establishing concepts and strategies of evaluation at the program operations level seldom has time to track down and study new developments in evaluation. Nor, does he have time to analyze how well the proposed concepts meet his own needs, to identify gaps, and to be in a position to help build more realistic approaches to evaluation.

This monograph is designed as a reference to contemporary evaluation approaches. It brings together a variety of old and new frameworks and ideas about evaluation and shows how they're related to one another. It identifies some of the literature on evaluation from a variety of fields and some of the people who are generating ideas about evaluation. It can serve you by:

1. Bringing you up to date on what has been happening to evaluation concepts.
2. Increasing your understanding of the present state of evaluative approaches.
3. Providing a basis for supporting or improving the evaluative approaches you're now using.
4. Bringing to your attention frameworks that can help you design evaluative activities.
5. Providing background when you need to discuss evaluation with consultants.
6. Identifying the areas where you need to be active and encourage others to contribute to further development of evaluation strategy and techniques that will be relevant and useful in evaluating your type of program.

The first section of the monograph provides an overview of the current situation. The last section presents brief summaries of contemporary models and approaches to evaluation. The middle section provides some guidelines for developing evaluation and a problem-focused index to the various models in the later section.

Those who want to get an overview of what's happening in evaluation will find the first section most useful. Those who want to improve or develop an evaluation strategy may want to skip the first section and go directly to Section II. However, some of the approaches included in the last section will seem out of order unless you have kept up on recent trends in evaluation literature or scanned Section I.

Section I

LOOKING AT THE CONTEMPORARY EVALUATION SCENE

This section presents an overview of what has been happening. It (1) identifies some of the influences that particularly affect evaluating programs for disadvantaged adults, (2) summarizes major developments in the conceptualization of evaluation that have implications for those programs, (3) introduces some of the types of approach that evaluation is currently taking, and (4) suggests the kind of work that still needs to be done before evaluation will attain its full effectiveness as a programming tool.

Needs and Issues in Evaluating Programs for Disadvantaged Adults

Heroic attempts have been made to evaluate programs for disadvantaged adults--those adults who are at a disadvantage, whether socially, culturally, economically, or physically in terms of the majority. Yet, there's often a general feeling of uneasiness and frustration associated with such evaluation because it has often proved less than satisfactory. There's a growing realization that something's needed that traditional evaluation approaches haven't been able to give.

Evaluation strategies are needed that work well within the environment surrounding programs for the disadvantaged. Although most of the needs and issues involved in evaluating such programs may also be inherent in other adult education programs, they're much more visible and create more pressure and frustration in new or controversial programs--especially programs with a broad base of involvement of governmental agencies and the public.

Examining some of the historical and environmental factors associated with some programs for the disadvantaged may help to clarify the kinds of evaluative activity needed and bring the causes of uneasiness and frustration into sharper focus.

Some of the influences that strongly affect evaluation of programs for the disadvantaged include:

1. Variance in viewpoints and needs.
2. Multiple units of programs.
3. Influence of state and federal funding.
4. Inherentness of conflict.
5. Lack of past operating history.
6. Relationship to established programs.
7. Assumptions about successful programs.

Variance in Viewpoints and Needs

One major source of frustration is the lack of understanding of the number of types of people who have needs for evaluative information about programs and the extent to which these needs differ. Although all adult education programs have a potential for five types of people--*local administrators, civic leaders and other influentials, administrators of funding agencies, program personnel, and program participants*--to be interested in evaluation, in most situations one or more of these groups are relatively passive. The interplay is between the needs of only one or two of the groups. However, in programs for disadvantaged adults, not only are they all apt to be active, but the needs of a sixth group, *social policy makers*, also exert pressure.

Social Policy Makers. Because programs for the disadvantaged are enmeshed in social policy and are instruments for carrying out certain

policies, evaluation is often requested and/or used by those trying to set or influence social policy--legislators who introduce and promote certain legislation, experts, politicians, leaders of causes. Those concerned with social policy want conclusive research data to defend or support policy decisions to show the value and effects of a program in relation to a particular social problem or policy.

Civic Leaders and Other Influentials. Some influentials have no direct connection with the funding of programs. Through their status position they do, however, influence prospective clientele and the kinds of support that programs receive. They're hard to identify. They may be bartenders, priests, street gang leaders, or cops on the beat. They evaluate programs and affect the program by their support, opposition, or indifference. Less is known about the criteria they apply or the kind of information they use. But because their position of influence usually rests in part on their being on top of things and giving good advice to their supporters, they have needs related to program evaluation.

Other influentials--those who serve in Congress, state legislators, or local governing bodies responsible for allocating public funds--want evaluation data as an input into funding decisions. Some are concerned only with the value their constituents attach to a program. Others scrutinize the extent to which the program produces the results it's supposed to.

Administrators of Funding Agencies. Administrators of the government or private organizations that allocate funds to programs not only need evaluation that helps them administer and manage their work, but also helps them provide the kind of evaluative information that the policy makers and other legislators (or the board of directors in the case of foundations) want. They're the men in the middle.

Local Administrators. Administration is valueless unless it's able to guide the unit to optimum fulfillment of mission. Local administrators need evaluative information not only to meet the requests of the funding agencies and as a basis for making internal decisions on budget, promotion, and salary increases, but also as a means of guiding staff in making basic program decisions. The administrator must take a direct and active role in program decisions. The administrator must take a direct and active role in program evaluation. He must use program evaluation as a base for providing administrative leadership. He also must be able to judge what evaluative strategies should have highest priority in helping his staff develop their own competency, in meeting the informational needs of outsiders, and in maintaining rapport with clientele.

Program Personnel. Although the programmer has real concerns about examining the results of his program, he must also use the evaluative processes that will give him the kind of immediate feedback he can use to carry out a program. Continuous evaluation during the programming process is as or more important than retrospective analysis.

While in many other adult education programs, programming is primarily a "one-man show" or a team operation of two or three professionals, programs for disadvantaged adults use a wide range of program personnel. Paraprofessionals are important. With their advent, the front-line professional has become partly an administrator and partly a direct programmer. In that role, he must not only be able to evaluate his own activities, but to help the paraprofessionals evaluate their work and draw them into a cooperative evaluation of the work of the entire team.

Too often program personnel are viewed as being one of the subjects of the evaluation or the provider of information. Often, too little attention is given to the kinds of program evaluation that will be of real use to them in programming.

Program Participants. The program participant receives even less recognition in evaluation schemes. His needs related to program evaluation are often bypassed. And yet, in programs for the disadvantaged, the evaluation made by the individual participant is of utmost importance. It determines his own activity. What he does--how he uses his energy, time, and limited money--are extremely important to him. He doesn't want to miss something that will be of value. On the other hand, he's unwilling to waste time on things that are neither relevant nor pleasant.

Sometimes the kind of evaluative information needed by the six groups coincide. Often it differs. For example, the kind of information the potential participant uses in deciding whether to take part in a program is different from the research type data that the social policy maker is striving for.

Not only does each of the six groups need evaluation information, but in turn, each may also need to know what evaluations the other groups have made of the program. For example, the programmer, program administrator, administrator of the funding agency, and social policy maker need to know what judgments the clientele have formed about the program. If it's important to assess participant reaction in traditional programs where the participants are usually motivated toward participation, and the programmer is experienced in identifying signs and symptoms of reactions, then it's doubly important that we find ways to assess participant reactions when the participants may not be as highly motivated and when the programmer may not be well versed in reading natural signs. Not only must a constant check be kept on how participants are reacting, but evidence of negative reaction must be fed back into the program immediately so adjustments can be made. The program unit is often more conscious of assessing the reactions of those to whom

they report, then in following participant reactions. This may be true even though negative reactions on the part of participants may result in no program participants and ultimately in no programs. However, which reactions to heed usually isn't an either-or situation. Mechanisms must be found for understanding the evaluative conclusions formed by everyone involved. There may be advantages in getting the interested parties in discussion that leads to the formation of mutual judgments.

Local program personnel--administrators and programmers--find themselves in the complex role of: (1) supplying information that others will use in their evaluations, (2) trying to get and respond to information about the kinds of judgments others are making, and (3) trying to get evaluative information on program process and results that helps them make operating decisions. What they do in relation to any of these three roles must be done in conjunction with or as time away from their other program responsibilities. They must be able to balance and reconcile evaluative activities within the parameters of the small amount of resources they can allocate to evaluation.

One way of reducing uneasiness, if not frustration, is clearly understanding the various needs involved and the amount and kind of activity needed to satisfy them. Then, better choices can be made about how the program unit will invest its evaluative resources.

Most models, frameworks, or approaches to evaluation have been designed for a specific type of situation. Such frameworks seldom handle the variety of other needs and evaluative responsibilities of local programming personnel. For example, a model designed to provide the social policy maker with sound generalizable data may have little value in helping the programmer guide and improve a program that is in-process, and vice versa. The kind of evaluative information provided by a particular evaluation approach needs to be examined

in terms of its value compared with other evaluative information provided by other approaches.

Not only must the six groups clarify and better communicate their needs and program administrators take a more active role in trying to establish balance among needs, but all six groups must take more responsibility for developing evaluation models. Most of the approaches included in this monograph--the ones that get the greatest attention in the literature--have been devised by outside observers, usually social scientists brought in to do evaluation. Though certain parts of evaluation can be contracted to outsiders, the users of the evaluation should maintain an active and dominant responsibility for determining what evaluation should do and what it should be.

Multiple Units of Programs

A program can be an activity carried out by one adult educator. It can be the total of his activities; it can be the total of what's carried out by everyone working for one agency or organization at one location. It can be the total of all like efforts carried on in several locations in the state or country. Evaluation is important in each concept of the program. But the evaluation frameworks differ.

For example, an approach useful in evaluating a local program may not fit the evaluation of a national program carried on in many locations. On the other hand, a framework developed for the national program may not give enough information to guide the local program unit.

As much as possible, relationships between evaluation policies and procedures at various levels should be clear and the criteria consistent. The actual processes and techniques may need to be specific to the nature of the programming unit involved. There needs to be further model building to fit each of the concepts of program. Most of our past models come either

from the national level or from the individual instructor's standpoint. Emerging models are more apt to focus on frameworks appropriate for a local program unit.

Influence of State and Federal Funding

Dependence on outside funds influences program evaluation several ways. First, of course, is the direct and concrete influence exerted by the stipulation that most federally funded programs must provide evaluative information. Unfortunately, the funding agency often fails to help the program develop a meaningful evaluation. Problems arise when:

1. Administrators of the funding agency (the men in the middle) don't know what they want and require evaluation hoping that the programmer or an outside contractor will know what they need and provide it.
2. The funding agency has specific types of evaluation it wants, but either is unable to communicate these specifics clearly or expects something completely unrealistic given the state of development of the field of evaluation and/or the budget restraints of the program.
3. The funding agency is willing to let the programmer choose the type of evaluation investment that will be most valuable to the program, but fails to either communicate this latitude or to stand by earlier communications. Situations are compounded when funding agency personnel change positions, responsibilities, or approaches to evaluation so rapidly that the person who okayed an evaluation plan has left by the time the evaluation is completed.

In some instances, record keeping and the collection of yearly "head counts" seem to be done for the sake of record keeping rather than for a useful purpose. If at one time there was a clear rationale for data requested in the myriad of report forms, the rationale has been lost or it hasn't been clearly communicated to those supplying the information. As a result, annual reports and statistics sometimes bear little or an unreliable connection to the evaluation of programs.

It's one thing to cope with evaluation when demands are clear and specific. It's another when the local level can't identify what's wanted and must expend energy trying to second-guess. In such instances, considerable

resources are spent on activities that have little value to either the program unit or to state or federal administrators. Often the programmer at the end of the chain doesn't understand what the Washington unit really needs. The message loses meaning as it filters down. Or, the Washington unit has lost touch with the realities of the situation and asks for information that doesn't adequately describe the operating programs and their results. Devising an evaluation system that's efficient and effective for all concerned in a multi-location, multi-level administrative network has different dimensions than a system developed to work effectively for the local program.

The need, both by those providing and those receiving funds, for a better understanding of evaluation and more accurate and efficient evaluation procedures triggered the expansion in ideas about evaluation that has occurred in the past few years. However, although progress has been made, much still has to be developed and decided. Federal bureaus will have to continue to be partners with theorists and practioners in further refining and developing evaluation concepts.

This discussion emphasizes the complexity of the evaluation needs, interrelationships, and importance of not trying to reduce evaluation to one specific procedure or to use only one evaluative approach. But federal funding also exerts other kinds of influences on evaluation. Influence comes with the nervousness among staff and clientele engendered by "soft" funds--funds appropriated for a set period only and then open to review and reappropriation. Nervousness runs particularly high with many programs for the disadvantaged because openness to understanding and political positions of leaders have been known to vary greatly over a few-year period.

Even initiating legislation is changed with later amendments. Programers faced with an urgent need to show the value of their programs also worry that political considerations will influence the funding, regardless of the presence or lack of evidence on the value of the program. Do you make a conscientious effort to evaluate or is it better to employ those resources in various forms of politicking? Certainly it's foolhardy to ignore politics in the broad sense of power bases and influences. Evaluation often has political involvements that must be assessed and dealt with.

Inherentness of Conflict

Seldom have adult education programs been so closely enmeshed in conflict as are programs for the disadvantaged. Conflict poses additional needs for evaluation and heightens the trauma involved with it. Conflict is either inherent or imminent in a variety of ways.

First, underlying conflict exists between the clientele and factions in the larger society: How representatives of both groups view the meaning of being disadvantaged, how each views and reacts to the suggested causes of that condition, and how each sees the responsibility of the other for erasing "disadvantage" present unceasing conflict that erupts at unexpected times and affects operating programs.

There's conflict in the use of tax monies. Some legislators, governmental officials, and groups within the public oppose the amount of money being spent on such programs. Others point out that the total budget is far less than the amount spent on defense or space exploration. Some say that if the federal money spent for educational and social action programs for disadvantaged clientele were given directly to the clientele, everyone would be above the poverty level. Others believe that a "dole" is incompatible with human dignity and investment in human development is well worth the added cost.

A third kind of conflict is the kind that sometimes exists between agencies who are competing for the same limited funds, for the attention of the same clientele, and for the same image and prestige in a community. Considerable disagreement exists about which programs are most valuable.

A program that hasn't found a foolproof road to success faces continuing conflict between programming philosophies and divergent program approaches and activities. For example, proponents of traditional class approaches, learning centers, and the problem-centered approaches may all vie for ascendancy in adult basic education.

Most of these conflicts involve valuing and reconciling or coping with opposing values. The programmer must be able to reconcile the expectations of the clientele with the expectations of the agency, with the interpretation of what the funded mission of the program is. Or, if unreconcilable, the programmer must choose from among the various expectations. At the same time, he must be able to successfully fend off competitors' assessments of the value of his program.

Evaluation can be an important tool in conflict situations. However, it must be a kind of evaluation that can deal with value questions. As yet this is a relatively undeveloped area in evaluation theory and in the systematic practice of evaluation. The scientist is trained to avoid valuing. Thus, his evaluation models may fail to adequately consider the process of dealing with values in evaluation.

Lack of Past Operating History

Many of the newer programs haven't been operating enough years to:

1. Create a generalized, positive image strong enough to provide a protective smoke screen and ward off demands for evidence of specific accomplishments.

2. Develop the internal protective mechanisms to handle, bypass, or submerge evaluation issues in the same way that established adult education programs have.
3. Use the wisdom of past experience in adjusting process and routine to improve programs. They're under the gun to achieve optimum operating effectiveness immediately and aren't given the same freedom of a span of years to grow and develop that other adult education programs have had.
4. Be able to draw on reservoir of research and experience germane to their endeavors; they must create as they go.
5. Be able to draw from a pool of experience personnel.
6. Create a positive, external image that attracts desired clientele.

When a program lacks past operating experience, it has a particular need for evaluation. For example, evaluation can help programmers build an understanding of their efforts that might otherwise take years to identify through ordinary activities. Evaluation can rapidly increase the information available on programming and reduce the need for trial-and-error activity. It can be used to build an image that ordinarily comes slowly through increased contact with the program by and through the participating clientele. When program personnel are alert and strong enough to prevent being shackled with old approaches to evaluation, lack of past history provides an opportunity to build evaluation strategies that meet their unique situation. However, making use of this freedom requires strength, perseverance, and the ability to develop logical rationale for the evaluative activities used.

Relationship to Established Programs

Another influence on the new program is its relationships to the old. Most programs either are appended to a traditional, well-established adult program agency or, although standing independently, come on the scene much later than other social service or education programs.

New programs are often ambivalent in relation to the old. When there are conflicts, should they try to maintain a solid footing with the established program? Or, should they relate to the needs of their clientele regardless of how the "relatives" react? Should they be concerned about meeting the evaluative criteria accepted by the parent agency or the old timer in their respective field? Or, should they concern themselves only with the criteria of the clients and sponsors? For example, should adult basic education programs be judged in terms of the standards applied to childhood education? Or, just as materials and approaches need changing to suit the new clientele, do standards also need to be different? Although new programs sometimes feel like "orphans," often an administrative or generic tie to an older generation exists that affects the new program's view of evaluation.

Assumptions About Successful Programs

Another very real but subtle influence on evaluation is the influence of technological progress and managerial science on attitudes about what it's possible for educational programs to do.

Because science can produce complex products in an orderly controlled process determined well in advance and carried through exactly as planned, it's assumed similar procedures always can and should be applied to human beings.

For example, we are apt to assume that:

1. It's always possible to plan with complete accuracy.
2. All people need the same things.
3. We're able to do anything we want to.
4. We can get conclusive proof of what happened.
5. We can always be efficient in our operations.
6. Technological activity substitutes adequately for human activity.

In some instances, these assumptions are well founded. However, in instances where they're not, the evaluation must be built in terms of reality rather than in false assumptions.

Our concept of the role of a plan has changed extensively over the years. It has changed from one of being a guide to that of being a contract. It's assumed, for example, that the programmer should always be able to identify in advance what results his program will produce and, like the factory assembly line, produce those results come what may. It's assumed that the programmer should by one means or another, be able to control the participants and other human environmental factors in the same way that man can control machines.

Many programs for disadvantaged clientele are producing greater and often more important side effects than the specific objectives of the program. One way of looking at the finding of important side effects is that the wrong objectives were set--the programmer couldn't diagnose what the real needs were. However, another view is that the exact effects of a given program input aren't readily predictable. It may have influence (helpful or harmful) far beyond its designed intent. Evaluation should consider those results as well as how the program accomplished what it said it was going to do.

Some evaluations become more an analysis of how well the program followed the original plan than an examination of how well the program met the original need. Traditional use of pre-post data collection in evaluation doesn't permit a program to change. Yet, the experiences of the past few years have shown that change is the byword. Few programs occur exactly as planned and produce the results that were originally identified. In such cases evaluation must be able to deal with questions like whether the change

was for the better or whether it was the easy way out. When you're evaluating a program rather than a plan, the plan should be treated as a means to the desired end, and the evaluation should emphasize whether that end was attained.

Mass production leads even knowledgeable adult educators who espouse philosophies of individual difference into the trap of expecting all participants to achieve the same things. Whether this is good or bad depends on whether those things are really important to all of the participants and the extent of varying needs within the group. When participants are extremely heterogenous in needs and interest, evaluation procedures should recognize and deal with individual differences. Success then would be determined by the number of people helped to do something important to them rather than by the average gain of the whole group.

Another influence business and technology fosters is a feeling of invincibility and unrealistic belief in our ability. Standards for success are often set in terms of the need or of a "pie-in-the-sky" hope of success. They assume that we not only can produce that pie, but produce it within one year. We either have too little accumulated wisdom to set realistic rather than idealistic standards for success or we ignore what we do know because we hope we can do more. We expect miracles--too few resources invested for too short a time to make mammoth strides in changing behavior. Goals that go beyond what we're sure we can accomplish do stimulate effort, but may result in unrealistic criticism when they're not completely attained.

Although efficiency is a quality to work toward, we may have little concept of what constitutes efficient education. With few exceptions, disadvantaged adults aren't an efficient clientele to help make major gains.

Although some are immediately able and highly motivated, many have suffered school and nonschool experiences that have reduced their faith in their ability to learn and interest in learning. The kind of input needed to establish trust and increase motivation is much greater than in many other programs. In addition, many of the clients face crisis interruptions that are more important than program participation. So, attendance is often irregular and dropouts frequent for reasons other than disliking the program. Evaluation is needed: (1) to determine what we can realistically expect to occur from programs with this clientele and (2) to find the best processes for being of greatest help.

As we transfer beliefs about success from the field of technology, we often forget all the years of work that went into the laboratory and field research that perfected the present assembly line procedures; we expect programs to produce the first time off the drawing board.

We may be misled, too, into believing evaluation can do everything we would like it to do. Just as our standards for program success may be unrealistic, so our standards for efficacy of evaluation may be equally unrealistic. Although we want conclusive and indisputable proof of accomplishments, how great is the chance of getting completely firm data when programs operate in a real and dynamic world where it's not possible to isolate participants and thus be able to study the effects of the program in isolation?

Making laboratory type research realistic in a living, uncontrollable world is one of the headaches of evaluation. In coping with that headache, however, we sometimes forget more important things; for example, the number of laboratory experiments that were made before data were accepted as proof of a scientific finding. If the resources needed for repetition of evaluation and the many analyses needed to develop even fairly conclusive proof were

actually expended, evaluation would cost far more than actual program operation. In addition, even though we amass results data on a given program, in how many instances do we have an adequate "season record" to be able to interpret the degree of success of a particular operation? Yet, we usually treat data not as input into the season record determining player averages, but as understandable and useable in isolation from comparative data on other programming operations.

In searching for objectivity, we may have been misled by the role of calculators and computers. Rather than considering the data as ways of helping human judgment be more objective, we sometimes think data constitute evaluation and accept, without careful scrutiny, evaluations with impressive data treated by highly sophisticated statistical processes. This transference means there's a tendency to define evaluation as the data involved in the activity rather than the conclusions of the evaluation or the process by which conclusions are reached.

We transfer some things from our technological successes, but we sometimes fail to transfer others. For example, the quality found when the product is evaluated at the end of production depends on careful evaluation of it at each step in the process. Yet some concepts of program evaluation ignore the importance of evaluating decisions throughout the total sequence of the program--for example, evaluating the purpose and objectives to see if they're realistic and attainable.

In some situations, most of our assumptions should be applied to programs. But, in others, they may be completely unrealistic. Concepts and strategies of evaluation must be founded on assumptions that are accurate within the environment of the evaluation.

These seven influences are only a few of the real-world conditions that affect evaluation of programs for disadvantaged adults and make it confusing and frustrating. The kind of evaluation models that are most needed are those that can deal with real-life situations and make valuable contributions in an everyday environment. Evaluation has to be realistic in terms of the real operational context of the program and its environment without becoming biased or distorted by real conditions. Recent experiences have caused evaluation to expand and change.

Contemporary Ideas About Program Evaluation

Ideas about evaluation are changing. Beliefs about program evaluation plateaued for a few years during the 1950s and 1960s when evaluation was equated with research methodology to such an extent that sometimes the terms measurement and evaluation were treated interchangeably. During that period, too, evaluation was often limited to determining whether content-specific objectives had been achieved.

Then the late 1960s brought an influx of new programs and new demands for evaluation. Established concepts didn't deliver. As a result, new ideas about evaluation emerged and new frameworks appeared. There's considerable divergence in those ideas. Most of them are still in the trial-and-testing stage. Many paths are being taken off the plateau of the earlier period, but few of those paths are widely accepted. None can be considered the main route. Some explore evaluation from the standpoint of its purpose, some from the standpoint of need, some from the view of organization and system, and some from the interactive elements involved. New definitions of evaluation are evolving (see Table 1).

Although the needs and pressures outlined earlier have influenced the changes occurring in evaluation approaches, development of theory and practice is not yet sufficient to completely deal with current needs.

Table 1

CONTEMPORARY DEFINITIONS OF EVALUATION

Evaluation is the process of ascertaining the decision areas of concern, selecting appropriate information, and collecting and analyzing information in order to report summary data useful to decision makers in selecting among alternatives.

Alkin, Marvin C. "Evaluation Theory Development." *Evaluation Comment*, II (October, 1969).

=====

Evaluation is the systematic process of judging the worth, desirability, effectiveness, or adequacy of something according to definite criteria and purposes. The judgment is based upon a careful comparison of observation data with criteria standards. Precise definitions of what is to be appraised, clearly stated purposes, specific standards for the criteria traits, accurate observations and measurements, and logical conclusions are the hallmarks of valid evaluation.

Harris, Wilbur. "The Nature and Functions of Educational Evaluations." *Peabody Journal of Education*, XLVI (September, 1968), 95.

=====

Evaluation is quality control of the processes and outcomes of an educational program.

Gottman, John M., and Robert E. Clasen. *Evaluation in Education: A Practitioner's Guide*. Itasca, Illinois: F. E. Peacock Publishers, Inc., 1972, p. 16.

=====

Evaluation is the determination (whether based on opinions, records, subjective or objective data) of the results (whether desirable or undesirable; transient or permanent; immediate or delayed) attained by some activity (whether a program or part of a program . . . , an ongoing or one-shot approach) designed to accomplish some valued goal or objective (whether ultimate, intermediate, or immediate, effort or performance, long or short range). This definition contains four key dimensions: (1) process--the "determination"; (2) criteria--the "results"; (3) stimulus--the "activity"; and (4) value--the "objective." The scientific method with its accompanying research techniques then provides the most promising means for "determining" the relationship of the "stimulus" to the "objective" in terms of measurable "criteria."

Suchman, Edward A. *Evaluative Research*. New York: Russell Sage Foundation, 1967, pp. 31-32.

=====

Social program evaluation is the systematic accumulation of facts for providing information about the achievement of program requisites and goals relative to efforts, effectiveness, and efficiency within any stage of program development. The factors of evaluation may be obtained through a variety of relatively systematic techniques, and they are incorporated into some designated system of values for making decisions about social program.

Tripodi, Tony, Phillip Fellin, and Irwin Epstein. *Social Program Evaluation*. Itasca, Illinois: F. E. Peacock Publishers, Inc., 1971, p. 12.

=====

However, the key ideas that are emerging can help you develop strategies that will work in your situation. Among the more valuable new ideas about evaluation are the following:

1. Program evaluation is a process rather than a procedure. It's generic rather than specific.
2. Program evaluation is more than examining the attainment of objectives.
3. Program evaluation is more than just evaluating the results of a program.
4. Program evaluation is more than instructional evaluation.
5. Program evaluation is different from evaluative research and program research.
6. Program evaluation is a management tool.
7. Program evaluation is people centered.

Program Evaluation Is a Process Rather Than a Procedure.
It Is Generic Rather Than Specific

The definitions in Table 1 are much broader than our old ones. They recognize that one specific procedural definition doesn't meet the range of needs or fully use the powerful potential of evaluation. Instead evaluation is most easily handled and has greatest utility if it's considered a generic term and used as such. Thus, program evaluation encompasses one or more generalizable processes, many evaluative strategies and approaches--more than 50 have emerged in the last 6 years, and specific evaluative activities.

The two most prevalent ways of looking at evaluation as a generalizable process are:

1. As a process of forming judgments about programs using criteria or standards of comparison and descriptions of what occurred and resulted in the program.
2. As a process of using information in comparing alternatives in reaching program decisions.

The first concept places the emphasis on judging and forming conclusions about program activities and program results. These conclusions are then fed

into decisions about further and future activity. The second concept emphasizes identifying alternatives and then using evaluation to help choose among those alternatives. The first concept would deal with a question like, "How important is this program?" The second would explore, "Which of these two programs is more important?" Or using another example, the first approach would examine how effective a given approach is in carrying out the agency's mission. The second would compare two or more approaches to see which is more effective.

There's a good deal of similarity between the two concepts of evaluation as a generalizable process. Both are so new that little attention has yet been given to how they relate to each other--how judgment and decision making function in relation to each other. Much more exploration of the linkages between the two is needed.

Accepting program evaluation as a generic term rather than restricting it to one uniform definition means that those requiring or proposing evaluation will have to be more exact in describing the purpose, processes, and procedures that they're referring to when they use the term "evaluation." Greater precision should mean greater clarity and more opportunity to fit strategy and approach to the existing environment.

Program Evaluation Is More Than Examining the Attainment of Objectives

A few years ago when evaluation was suggested, we automatically brought out a framework that examined the extent to which objectives had been met. This, of course, is a fairly important evaluation to make of a program. But, it's not the sole, and sometimes may not even be the most important, analysis of results. Two of the major emerging developments contributing to a better understanding of evaluating results are: (1) the expansion of

what needs to be assessed when you're concerned with examining the results of a program and (2) the distinction between describing and evaluating results.

More Than Objectives. Evaluation is beginning to look beyond objectives. The growing number of results approaches that go beyond the results stated in the objectives aren't ignoring objectives nor discounting their role in guiding programs and bringing results into better focus. They are, however, recognizing that in the complexities that exist in most programming situations, programs are apt to produce unanticipated results. Sometimes the values or harms of these unanticipated results are more important than the objectives per se. For example, with some clientele, the fact that they're in social contact with others in the program may have value far and beyond what they actually learn from the program. In some instances, programs that show few results in terms of their content-specific objectives may still be making extremely valuable contributions to the participants and society.

Harmful results also need to be considered. For example, a person may learn the multiplication tables, but learn to hate arithmetic at the same time. A person may meet the objectives of a program very well, but be so disconcerted by it that he won't return to it or similar educational experiences.

Many of the broader approaches compare the amount of results produced to: (1) the needs that initiated the program, (2) the kind and amount of results that must be produced if the agency is to attain its mission, (3) broad competency standards, or (4) statements of the kind of results that a program conducted with the particular clientele using a certain amount of input should be expected to produce. Objectives are symbols or transmittal links to one or more of these broader definitions of results such as meeting a need, or carrying out a mission. If objectives are well chosen

and directly on target, results judged against the broader base will also show that the objectives have been attained. However, if the objectives aren't germane, their attainment doesn't necessarily mean that any of the four broader bases have been achieved. In many instances, objectives become so narrow or so side-tracked that little is accomplished through them. An evaluation issue is: Should programs only be examined in terms of what the programmer thinks he's doing (that is, what he has defined as his objectives)? Or should they be examined in terms of what he should be doing (that is, what should be happening if needs are being met, mission accomplished, broad competency standards attained, regardless of what his objectives were)?

Evaluation that's concerned with the overall effectiveness of a program is concerned not only with results in terms of behavioral changes in people, but also with the proportion of the potential clientele that's reached, the balance in types of people reached, the extent to which the results deal with urgent and continual need, and the care with which participant, agency, and societal resources are used. Program evaluation is as concerned about the value and suitability of the program as it is with whether its purpose is accomplished.

Description Vs. Evaluation. The second major step in improving understanding of what's being done when a program is evaluated in terms of its results is the need to distinguish between *describing* the results of a program--providing evidence of what results did occur--and *evaluating* those results. Past approaches to result evaluation limited themselves primarily to identifying if results did occur and testing whether they probably came about as a result of the program. Seldom did they try to identify the sufficiency of those results. A statistically significant chi-square value in a comparison between program participants and a control group was considered enough, even though half of the participants hadn't achieved the expected results.

Evaluating results gets into questions requiring more information than just whether the program produced results. It involves such sticky questions as: Are the results consistent with what's logical to expect? Are the results important? Do they contribute more to the participants and society than if the time and other resources had been invested in other things? Were they produced at a reasonable cost? Are the results sufficient in terms of the overall need? Are they sufficient in the expectations of the participants and the amount of time and energy they invested? Is there any evidence that it's realistic to expect a program to produce more results than this one has, given the same budget, personnel, and working conditions? If the results are insufficient, does it mean that the program isn't effective or that changes need to be made in the way the program's carried out? Have the results been replicated with enough consistent findings to clearly decide on the value of this particular program? Are the results of one program significantly more valuable than those that could be produced through an alternative program with the same expenditure of resources?

Most of these questions can't be answered by traditional evaluation approaches. Ways of dealing with these types of questions still must be developed.

Program Evaluation Is More Than Evaluating the Results of a Program

One of the most important of the recent developments is the distinguishing of the evaluation of program results from the evaluation of programs. For years, evaluation has been thought of as primarily a summative activity in which the results of a program were studied. Now formative evaluation--evaluation that influences that program while it's progressing--is gaining

equal importance. Evaluation of each of the key processes involved in programming (planning, management, interaction, etc.) and of each of the key stages or activities is essential to achieving results.

Evaluating the extent of need for various types of programs and using evaluation in setting priorities is imperative. Evaluating the outcomes expected of a program to see if they're realistic and focus on the most crucial needs, evaluating the plans for the program to see if adequate inputs are being marshalled and if the activities planned can produce the type of results expected, and monitoring the processes involved to see that they stay on target and that repetition or reinforcement are added to the original plan if such activity is necessary--all are essential if a program is to achieve maximum results. Evaluation that influences ongoing developments in the program has great value--it improves and gives immediate benefits.

More is involved here than the usual categorization in terms of process versus product evaluation or as some people say, the evaluation of means and the evaluation of attainment of ends. Program evaluation as a generic term can include the evaluative activities that focus on either process or product. But the emerging emphasis is on frameworks that consider both process and product and, more importantly, the interrelatedness of the two. Program evaluation deals with the program as a functioning, producing system. It includes results as one key component, but also examines other aspects of the program. Processes and structural components such as teacher's performance, materials, and facilities are examined in relation to how they can be more effective in generating a viable, valuable program. The emerging models differ in how they define the parts of the functioning system. Some program evaluation approaches focus on processes, others on program components.

Program evaluation serves both during the actual program operation and in retrospectively analyzing how the particular level of results was produced. At a broader level, it expands our knowledge of what makes up successful programs under specific environmental conditions (backgrounds of students, extent of budgetary support, social pressures). Such models are particularly important to new and/or controversial programs.

This broader track of program evaluation emphasizes developing and improving programs. Approaches that focus only on results are primarily concerned with questions about the termination or continuance of the program. Both are essential. One doesn't substitute for the other. Every program should be reviewed periodically to see if it should be continued. (Old programs never die. They don't even fade away.) But when new programs are developing and when decisions have been made that established programs will continue, program evaluation that emphasizes improvement is crucial.

Program Evaluation Is More Than Instructional Evaluation

Until recently, "program" has been a little used word in educational circles. Few school personnel felt a need for any terms beyond curriculum and instruction. The Tyler approach in the 1950s clarified the relationship between evaluating students and evaluating instruction. It used objectives as the focal point to guide the data gathering on student performance.

Adult education began equating the term program evaluation with the Tyler construct in general education that was primarily concerned with the evaluation of instruction. Thus, for some time program evaluation was seen as being how well participants achieved the instructional objectives set for the group. Now the distinctions between evaluation of instruction and program evaluation are becoming clearer.

Instructional evaluation is concerned primarily with the specific course or program activity. Program evaluation is concerned with the additive effects of a series of instructional components. What are the combined results of the courses taken in a two-year apprentice program? Instructional evaluation is usually most concerned with knowledge, skills, and attitude change. Program evaluation is concerned with the impact that those changes and program participation have on the person and those he's in contact with. Instructional evaluation is more apt to deal with how the program satisfies the specific needs of individual learners; program evaluation is more apt to deal with how the program meets the needs of the community or a subsection of society.

Program evaluation includes instructional evaluation, but deals with additional things. It's concerned, for example, with establishing priorities among instructional programs. It's concerned with the extent to which the instruction is attaining the institutional goals of the agency. As an example, if the agency (institution) holds a broad objective of helping adults develop ability to think analytically and creatively, how well is such an objective being accomplished within a specific course or program activity?

Program evaluation is concerned with whether adequate resources are being applied to the right programs. It deals with such issues as whether to be all things to all people, or whether to concentrate resources in certain priority thrusts. Instructional evaluation is concerned with the quality and results of specific activities within the total program. Program evaluation is concerned with the extent to which the agency is carrying out its mission. Instructional evaluation controls and improves the specific programmer-participant relationship.

Program evaluation doesn't substitute for instructional evaluation; and instructional evaluation doesn't substitute for program evaluation. Even though good instructional evaluation provides essential information for evaluating

a program, it doesn't provide all of the information that's necessary even in programs where instruction makes up the total of the program. It's possible to do a good job of instructional evaluation without ever really doing program evaluation.

Program Evaluation Is Different from Evaluative Research and Program Research

Considerable debate is currently going on about whether evaluation and research are the same thing. There seems to be three related activities that use similar methods. All are important in programming.

Program research looks for new and generalizable knowledge. It expands our understanding of programming. Findings are useful in a variety of settings. It's trying to formulate theories and principles. Occasionally program evaluation provides a context for program research or program research uses program evaluation as one of its parts. In these cases, however, benefits to the program being evaluated are secondary. The focus is on an audience way beyond those involved in the program used for the research.

Evaluative research is a specialized branch of research that deals with valuing. It's concerned with identifying the value of programs or with comparing the values of two or more programs. It's concerned with whether the type of program can produce something of value in relation to a need. It differs from program evaluation in that, as a form of research, it's interested in these questions only in the broad context of generalizable answers, answers that contribute new knowledge and lead to theory and principles. The contribution to the particular program examined is secondary. Evaluative research may be an important procedure in program evaluation at a national level or in new programming situations.

Program evaluation, as distinguished from research, usually deals with old questions about programs. It's not trying to get generalizable information--it's

trying to get sound and reliable information to use within the context of the specific program. Program evaluation is a field in its own right. It resembles its research cousins, but there's a growing awareness that it shouldn't be treated as a twin. Its processes and procedures need to be developed within its own context rather than borrowed directly from other parts of the family.

The confusion of evaluation and research has led to underdevelopment of evaluation procedures in at least five main areas: suiting data techniques to evaluative needs, understanding criteria, distinguishing between description and judgment, providing useable information, and communicating evaluative findings.

First, although evaluation may borrow data treatment methods like precise instrumentation and statistical processes from research, they must be tested for validity and reliability as tools in evaluation rather than automatically transferred and applied as they'd be in research situations. Considerable adaptation may be needed. For example, when is it best to examine average gain for a whole group and when is it better to examine the number of students making specific amounts of gain? Is it necessary to rule out chance occurrence at the .05 level, or would a .40 or .25 level be adequate in evaluation? Much of test development is based on procedures for discriminating among students to assign grades in line with a normal curve. In evaluation, such discrimination is irrelevant. The type of data recorded on student-teacher interactions might take a different form, when a program is being evaluated than it does when it will be used to test a particular research hypothesis. A good many other examples could be given.

Although we're becoming aware that new procedures are needed and standard research techniques may need to be amended and adapted, this is a relatively new field. Some evaluators and researchers, for example, have been suiting the

statistical significance level--the level at which chance occurrence is ruled out--to the context of the problem for some years now. But little has been done to identify the conditions under which a .25 or a .10 level may be appropriate in program evaluation. In many instances, evaluation may not require the same degree of stringency in data dealing as does research. This may be particularly true when the consequences of inaccuracy are relatively small.

Second, criteria must receive much more attention in evaluation than in research. In research, hypotheses serve as the basis for deciding what information to gather and how it's to be organized to be interpreted. In evaluation, criteria--decisions about what the program will be judged on--play the same role. Evaluative criteria aren't always the same ones that apply in research, although there's a similarity in the role of criteria in the two situations. Unfortunately, most research texts assume criteria without dealing with them as important phenomena. The fledgling researcher has often been so enthralled with the data-handling procedures that he hasn't mastered the concept of criteria as it applies in research. So, he's unable to carry even the basic concept into evaluation.

Although some program criteria presently exist either in a formal or subconscious state, they must be probed for clarity, reality, and the extent to which they portray the most valuable things in programs and programming. Programs can appropriately be judged on a wide range of criteria, some of which are in conflict with others. Criteria need to be ordered in terms of their importance. For example, which program is more successful--the one that reaches a large number of people with a moderate amount of help to them, or the one that gives great help to a very small number? Which program is more successful--the one producing efficient results for a short period of time, or the one with only moderate results with a longer retention time?

Criteria describe what's of value or good. They're debatable. Although programmers sometimes wish to avoid the trauma of debate, a confrontation on criteria can clear up some of the ambiguities and frustrations that plague evaluation.

Program personnel and their professional groups are responsible for developing, screening, and debating criteria. In doing so, considerable attention should be given to the views of clientele. Although experts knowledgeable in a given field can make recommendations and researchers can help clarify issues and determine the degree to which certain criteria are commonly held, it's not the outsider's prerogative to establish the criteria for the evaluation.

A third major area for development is the greater understanding of the difference between description and evaluation. A presentation of research type findings about a program usually constitutes a description of the program or of its results. Evaluation doesn't really occur unless judgments are made about such things as whether the processes used were suitable and of high quality or whether the results produced were enough for the resources expended, or how effective the processes were in producing the results. Research tries hard not to be judgmental. Evaluation, on the other hand, doesn't fulfill its role unless it takes responsibility for delivering conclusions that do judge, identify the adequacy of what's being described, or show relative advantages or disadvantages of two or more alternatives.

It's probable that in some situations, particularly when reporting to outsiders, the programmer may only want to describe what happened in the program and/or its results. However, in decision making, descriptions are useless unless judgments are made and values assigned. Unfortunately, sometimes adult educators who have majored in science fields and had graduate

courses in research methodology not only haven't had experiences that help them improve their ability to judge programs, but also have been taught to mistrust and avoid situations where value judgments are involved.

A fourth major area for development deals with the importance of being able to use data. It's not enough to just collect accurate data. Those data must be accurately and immediately interpreted within a specific context. Interpretation rests with man. Wisdom and experience in forming and using criteria, in assessing the limitations of data and the potential consequence those limitations pose, and finally in making, communicating, and defending judgments are more important skills in evaluation than are the skills of data gathering and statistical analysis. Scientifically produced data are a valuable input in evaluation, but seldom should stand alone as the output from it.

From one perspective, data-gathering activities are a separate function from evaluation. They're essential to it, but can be considered as prior and prerequisite to the actual act of evaluation. Unfortunately, as we've borrowed from research, we've been more apt to focus on the data-gathering techniques than we have on the part of research dealing with using data. We've hoped that having high quality data would solve our problems. We're finding that data in and of themselves give few answers. The next step in understanding evaluation and its procedural development must deal with processes of the mind that are essential in using information in evaluating programs.

Fifth, our associating evaluation with research sometimes leads to unproductive ways of communicating results. We may be too prone to prepare reports in traditional research ways that relegate conclusions and implications to the end of the report. The people who should use evaluation seldom get to the end of the reports. If they do, they are often so

confused by academic hedging and research jargon that they can't decipher what the evaluation really means in terms of their own needs and situation. Currently a variety of ways of inserting evaluation findings and conclusions into the bloodstream of the program and community are being emphasized.

Meaningful evaluation has been stunted by failure to recognize that program evaluation and evaluative research, although related, aren't the same. Now that the nature of the differences are clearer, progress should be made in developing specific procedures that work well in evaluation.

Program Evaluation Is a Management Tool

Evaluation can be a powerful working tool in programming. It's not an end in itself, something to engage in because it's intrinsically good, but a way to get things done. With the clearer understanding of evaluation's role in decision making, comes a better understanding of the value of evaluation in guiding and managing program activities. Evaluation provides a basis for better program choices and for more rapid response to needs for improvement. It can be a tool in improving total operating efficiency as well as providing the client and society with more effective programs. Distinguishing between evaluation and research and clarifying the role of research-type data in evaluation, paves the way for a clearer understanding of evaluation as a natural human process--a continuous programming activity rather than an episodic or extensive but infrequent effort. The emphasis then is on improving accuracy of this process by introducing more system into the way it's carried out rather than by replacing it completely with periodic, systematic, formal evaluations. Evaluation in terms of intuitive judgments does and should go on continuously. It should be an automatic part of an adult educator's professional skills. He should form the basis for making accurate judgments as a part of gaining professional competence, in

the same way a doctor makes accurate diagnoses. Just as the doctor uses lab tests when he feels they're warranted, so the adult educator should know when more extensive, systematic processes are needed to improve the soundness of his judgment and when systematic processes are worth their cost. Both natural and systematic evaluation play important roles; neither replaces the other. Both are essential in managing programs.

Evaluation has to be well managed if it's to serve as a management tool. The degree of time invested in evaluation, degree of formality, and objectivity of procedures should be matched with the degree of value the particular exploration can produce. Systematic evaluation is a costly activity. The degree of system used should be equivalent to the importance of the results produced. Some situations require the best possible criteria and most scientifically pure data it's possible to get, regardless of the cost in time, money, or inconvenience to participants. In other instances, following the usual systematic procedures will require too great an investment for the importance of the judgments to be made. Attention must also be given to what information needs to be instantly available in an up-to-date form. The typical point in time, or episodic, research-like effort often doesn't produce results in time to influence ongoing decisions. One of the challenges in designing the information systems that provide input into continuous evaluation and decision making is reconciling the cost of keeping information continuously up to date with the extensiveness and value of its use.

The overall evaluation strategy of an agency or programmer should give equal attention to the everyday processes of evaluation, "diagnostic" evaluation, and more systematic endeavors as required. Evaluation should be designed for the context in which it will be used. What questions must be answered or decisions made? By whom? By what date? A clear understanding

of why you're evaluating and what you want to accomplish by that evaluation is essential in effectively using evaluation as a management tool.

Program Evaluation Is People Centered

Another major development in evaluation is the increased emphasis on evaluation's relationships to people. Evaluation as input into decision making emphasizes the need of interface and interaction. Evaluation for program improvement recognizes that those who must make the improvements must be actively involved in the evaluation. Recognition of close relationships between evaluation and politics and policy emphasizes the human element. The growing awareness of the central place of criteria and judgment's in evaluation emphasizes that those criteria come from people and that judgment is made by people. Even in terms of data, the need for involving a variety of people in interpreting data so that a more complete picture can be secured is becoming more apparent.

For a few years we tried to take evaluation out of the people realm, hoping to increase its objectivity. Lately, however, there's a growing return to active involvement of many minds as conclusions are formulated and decisions made. We need ways of increasing objectivity while relating closely to the people involved during the evaluation process.

More responsibility for evaluation is resting with program personnel. The role of outside evaluation consultants is changing. In addition to helping the system cope with data, the consultant must be able to serve as a catalyst and help program personnel understand and use evaluation findings. He must be able to help the system design evaluation that will be useful and meaningful to it. This may be different from the traditional assignment of designing a research approach to examine results.

Perhaps its important to caution you that you've just been riding the vanguard of the thinking about evaluation. These seven ideas have not been widely accepted yet. However, they hold promise for making evaluation a much more valuable force in programming.

The discussion of contemporary ideas has been general and hasn't related specifically to programs for disadvantaged adults. What do these changes mean in terms of such programs? Here are a few implications.

1. Pressure for evidence of results is great. However, even though much energy is going into this type of evaluation, equal attention should be paid to evaluation that helps programs develop and improve. Many programs for disadvantaged adults are really in very preliminary stages of development--the search and error procedural refinement stage--where good formative evaluation is crucial.
2. A second look should be taken at the traditional ways of using and interpreting research type data on results. Are the standards being used realistic or are they laboratory standards and not life geared?
3. Those programs where past experience is limited or controversy is inherent may want to be sure to examine a wider range of results than just those specified in the objectives.
4. At least two types of evaluation are needed in most local programs for the disadvantaged. One monitors instruction (or other type of programmer-participant interaction) to provide feedback to the participant and to help the programmer make decisions as the interaction continues. The other deals with broader program questions, such as are the right people being reached and the most efficient approach being used?
5. Because of the urgent need for becoming effective and efficient within a rapid time span, evaluation that facilitates better program management is extremely important in programs for disadvantaged.
6. A people-centered approach to evaluation that involves both staff and clientele in determining what will be evaluated, what criteria will be applied, how data will be interpreted, and in forming judgments and recommendations is particularly important in programs for the disadvantaged.
7. Finally, perhaps the most important implication is that which comes from accepting evaluation as a generic term including a range of strategies and approaches. Programs for the disadvantaged will get the most mileage out of evaluation when the approach is suited to the need and the programmer or agency has a well-stocked kit of evaluation tools.

We've now set the context for examining some of these evaluative tools.

Current Approaches to Evaluation

The last section of this monograph summarizes more than 50 approaches to evaluation. The terms "approach," "model," and "framework" are used interchangeably. Some are preliminary sketches; others are extensively developed. Some are adapted from processes that were used; others are designed as general ways of conceptualizing evaluation. Few of the authors would consider their materials "models" in either the theoretical or the "ideal" sense, yet all provide ideas to the programmer in helping him set evaluation strategy.

Even though this monograph includes a sizeable number of summaries from a variety of social fields, it isn't exhaustive. New models are constantly emerging. Yet, there's need for considerably more work--particularly on the part of those vitally affected by the evaluation.

The reader that finds the number appalling will probably ask: "Why so many?" "How do they relate to each other?" "How can I sort them out?" "Which should I pay attention to?"

We'll talk for a moment about the first two questions. Section II is designed to tackle the third and fourth.

In part, the proliferation of approaches represents a groping and a search for relevant models that was triggered when the established approaches to evaluation were found wanting. The number of approaches may decrease as acceptance and new stabilization occurs. On the other hand, the proliferation also represents the awareness that many kinds of evaluation are needed.

The situation can be likened to examining a mountain. The geographer is interested in topography. The geologist in rock samples. The mountain climber, the engineer designing a railroad, the pilot flying over the mountains, and the native going from one valley to another are all getting

data about the mountain range to try and describe, evaluate and deal with it. But, the data gotten and the approach to getting that information are far from the same.

In many ways programs are like a mountain. They're complex with many planes and facets. These planes and facets can be examined in a variety of ways. People have varying needs and purposes in relation to describing and evaluating them. Many types of information exist that are useful to those purposes. Therefore, just as there are many ways in which a mountain can be described, analyzed, and evaluated, so are there many ways that programs can be examined and evaluated. The approach you take depends on your particular needs.

For centuries, people thought they knew what mountains were. Then art and science provided new tools of perception and analysis and our understanding increased many-fold. We're moving toward that stage in program evaluation, but as yet it's more a random, searching exploration than one clearly governed by distinctly different approaches. We're at the pretaxonomic stage. Little has been done in categorizing approaches. But from very rough preliminary classifications like that developed in this section, you're better able to identify central routes of approach and gaps.

The analogy leaves us vulnerable to the questions, "Are we making mountains out of molehills? Must program evaluation be this complex?" Perhaps one of our problems has been that we've tried to treat programs like molehills and to use only one analytic approach to them. By recognizing their complex and dynamic topography and by accepting evaluation as a complex set of tools from which appropriate ones are selected to fit the aspect that is being examined, we may actually make evaluation simpler and less frustrating.

But why include so many approaches in a reference designed for the practitioner? Several alternatives emerged. (1) We could limit the monograph to the approaches that are most fully developed as models in the academic sense; (2) select prototypes or examples of different forms of program evaluation and let them serve as representative of the others; (3) narrow the definition so that only one subgroup of approaches is included--for example, only those drawn from education or even more specifically those developed for adult education and programs for the disadvantaged; or (4) present as full an assortment as possible, but provide some general type of grouping for ease in sorting and selecting approaches.

We chose the latter alternative for several reasons. First, few of the frameworks are sufficiently developed and tested to stand out as models for the field.

Second, although similarities exist, each approach has at least one unique idea that sets it apart from its cousins. The summation of these ideas may be of more value than any one approach as a separate entity.

Third, although the materials developed for adult education are extremely useful, they don't as yet provide as wide a set of choices as do other fields. Although sources could have been limited to education, some adult education programs for the disadvantaged are much broader than programs developed within formal education and are more like those of health and social services. Therefore, a few approaches from other social fields have been included.

Fourth, including a fuller range increases your perspective of program evaluation as a generic activity and your understanding of the range of approaches that are useful.

Fifth, you may be interested in how the models and approaches you're familiar with relate to one another and to the broad field.

Finally, and most important, there's no guarantee that one approach will be more useful to you than another. Ten readers are apt to be caught by 10 different approaches. What has value in one situation may have little value in another.

The next question is: If we're not restricting the number, how do we organize them? The mind can't retain and deal with more than a few items at a time. Therefore, how can the various approaches be sorted and classified to help you grasp them? Again several alternatives appeared: (1) sorting in a technical, taxonomic form; (2) sorting by source of origin; (3) sorting by extent to which approaches are commonly known or (4) developing some general groupings that may be useful to the practitioner. We chose the last alternative. However, preliminary efforts in regard to the first three systems of organization are included in the appendix if you're particularly interested in classifying evaluation approaches or want to use source of origin (which ones were developed for adult education programs) or degree of familiarity.

We decided to try to form a few general groups based on some aspect they had in common that set them apart from other models.

First, the approaches that *pattern a functioning system* were separated from those that focus on *program results*. The approaches that pattern a system were then divided into two categories--one dealing with patterning of programs; the other with patterning of evaluation. Each of these two categories was again subdivided. The program patterns were divided into those that emphasize evaluation as input into decision making and those that pattern parts of programs. The approaches that emphasize evaluative activities were divided into those that deal with kinds of data and those that describe processes of evaluation. Then the multitude of approaches that focus on program results were separated into those dealing with objectives and those dealing with a broader approach to outcomes and effects.

The following six groupings resulted:

1. Evaluation as input into decision making.
2. Evaluation of program parts.
3. Evaluation--kinds of data; types of activities.
4. Evaluation processes.
5. Results--attainment of objectives.
6. Results--evaluation of outcomes and effects.

The grouping provides one way of categorizing current approaches to evaluation. Table 2, page 50, gives the location of various approaches within the 6 groups. Although several of the approaches could have been placed in more than one category, we grouped them according to their primary focus.

The various approaches don't substitute for each other. Few are in conflict with each other. Many can be used in combinations of two or more. Even the approaches within a group aren't different ways of doing the same thing. For example, you could and perhaps should use all of the approaches in Group 1 over the period of a year. The *decision centered evaluation model* (CIPP) provides a framework for dealing with the decisions necessary in programming over a span of time. Although it can be applied to a specific project or program phase in the same way that the *developmental evaluation approaches* are, it has additional value if applied to the total program operations. The *differential evaluation model* emphasizes judgments and conclusions more than decisions, but provides an approach to deciding what aspects of the program will be judged. It's useful in taking stock of what's happened or what's happening at a given point in time. The *discrepancy evaluation model* is useful in trouble shooting. If a program is ailing or if you don't really understand what's happening, the discrepancy approach is helpful. The *priority decisions approach* is a subpart of context evaluation. But because the ultimate worth of a program depends on the choices made at the

very beginning, frameworks for evaluating the fundamental decisions of a program (which clientele, what content, how long, what approaches, etc.) are extremely important. Finally, most programs use texts, films, or other materials and need ways of choosing among them.

The second group of models can be used in conjunction with the first with no difficulty. The first group deals with qualities or stages of programming. The second deals with actual parts of that program. You could, for example, combine the *actual component approach* with the *decision centered*, *differential*, or *discrepancy* evaluation approach. The *actual component approach* would identify what elements within the program would be studied. The approach selected from Group 1 would establish the program sequence of examination and reexamination. The approach selected from Group 2 would indicate what element would be examined.

The third group emphasizes parts much as the second does, but it's talking about kinds of data that are used rather than about the program part directly. Similarities and differences exist among the approaches in Group 3. For example, the *countenance of evaluation* and *system role model* were developed during the same period by men at the same university both applying system and role theory. The *means-ends hierarchy* provides a different way of looking at data from the first two. It limits itself to the process and product categories of the *decision centered evaluation model*, but distinguishes various categories of product data.

The fourth group of approaches deal with processes. Three of them are basic processes that can be applied regardless of what it is about a program that's being evaluated. The two traditional processes, *appraisal* that emphasizes human judgment and approaches that emphasize a high use of *systematically collected data*, aren't alternatives to each other. The *natural process approach* indicates that they can and should be used together.

The other three approaches deal with specific applications of process. The *monitoring evaluation* model stresses the continuous approach to evaluation using either appraisal or systematic data processing or both. The *improvement evaluation* model shows how appraisal can be used interactively to stimulate programmers and the *transactional evaluation* emphasizes an approach that will help staff accept recommendations for change.

The fifth group is the most highly developed. Much more attention has been given to the details related to evaluating results using *attainment of objectives* as the basic criteria of success than to any of the other forms of evaluation. Rather than emphasizing a few of the several developments, this section gives a fairly concentrated overview of what has happened related to objectives.

The sixth group again gives more space to the individual approaches. Most of the approaches to results that are broader than objectives try to define and describe the types of *outcomes and effects* that make up program results.

Groups 5 and 6 can be subsumed within the approaches in Groups 1, 2, and 3. They're ways of spelling out the results component in more detail when you're concentrating on taking a program system approach to evaluation, with results being one component in that system. On the other hand, if you're particularly concerned with results and trying to explain the degree of success in producing results, some of the frameworks from Groups 1, 2, and 3 will be useful in addition to those within Groups 5 and 6 that try to use program processes or elements as explanatory variables. Although it's unwise to separate results from other program elements in that you lose much in understanding how a program produces or fails to produce results, sometimes outside pressures for results data and lack of resources force you

to limit exploration only to examining the results of a program. Approaches that emphasize results have been left as separate categories, but have been included as part of program evaluation rather than defining it as a kind of evaluation important in its own right, but different from program evaluation.

At this point, unless you're already familiar with some of the approaches, the details in the last few paragraphs that have discussed the various groups of approaches may have little meaning. Don't worry about it. The intent of the paragraphs is to underscore the idea that the various approaches to evaluation seldom deal with the same need. The search shouldn't be for finding a way, but for building a repertoire of ways of evaluating.

In this section and in the summaries in Section III, the approaches aren't specifically discussed in terms of evaluating programs for the disadvantaged. Such programs range extensively in terms of their nature and their needs. The range is as great as in the whole of adult education. Therefore, the application of the approaches and the potential for using them isn't unique in this particular kind of programming.

The Work That Lies Ahead

Although you'll find this monograph very useful, it may not meet all of your needs. Its value is limited both by the way the monograph is structured and by the present state of the theory, art, and practice of evaluation.

Limitations within the monograph include:

1. The nature of the classification system. It's tentative and general. It may be misleading.
2. Omission and distortion that comes through summarization. Some of the original sources are books, others are brief papers. In some instances, we've heard the author present the approach; in others we've relied only on written source materials. It's difficult to summarize and make the original ideas clear.

3. The lack of parallelism. Even though a standard format is used in presenting the approaches, it's very difficult to compare them. They aren't at the same level or stage of development. There's no consistency in overlap and duplication. It's there, but not sufficiently to make a clear pattern.
4. The lack of clear procedural directions. In summarizing the approaches, priority has gone to the parts that outline the "what" and "why" of evaluation rather than those that deal with the "how" of carrying out procedure. Once you have a clear sense of direction and specification, it's fairly easy to generate a process that will aim in the direction you want to go. Some of the sources give procedural details. Others don't.
5. The lack of case examples. In some cases, the models haven't been completely applied yet. In others, it requires too many pages to do justice to a description of the application. In some instances, a particular framework can be applied in several ways. Just presenting one way restricts you too much. Unfortunately most evaluation reports are designed to give information about programs, not to describe the evaluation. At best they stop with conclusions and recommendations. Few give the full story as to purpose, rationale, overall approach (procedure usually deals only with how data were handled), and what happened because of the evaluation. To understand the effectiveness of a particular evaluation approach, you have to know what occurred as a result of it.

Various limitations in terms of the present state of theory, art, and practice in program evaluation have been indicated in the earlier discussions. A good deal more needs to be done in conceptualizing modeling and developing and testing guidelines and procedures. For example:

1. Now that we've realized that accurate information is an input into evaluation but isn't the whole of evaluation, much more attention must be given to translating the mental and social interactive activities of evaluation into procedures or guidelines. How do you assign value? Control trade-offs? Exactly how do you go about assigning meaning to data and forming conclusions? How do you evaluate the differences between alternatives and finally reach a decision? How do you develop a logical argument to support a judgment that's powerful enough to convince an antagonist? These answers aren't easily available in procedural steps and yet these are some of the activities that constitute the heart of evaluation. We need frameworks for making the mental activities of evaluation more systematic and accurate. Some of the procedures must be program specific. They need to be developed by a team representing all those who have interest in the evaluation--participant, programmer, administrator, legislator, social policy maker, local influential--and designed to fit within a certain programming situation. Professional groups must take a more prominent role in developing standards and setting realistic input-output ratios.

2. We need guides and procedures for helping the professional improve his ongoing, everyday evaluation and decision making. How can he sharpen his powers of accurate observation and interpretation of what he sees and hears occurring? How can he detect biases and the possible effects of those biases?
3. We need cost-benefit studies of the contribution of "hard" data--that which is gathered using research type techniques--to accurate evaluation. Are the judgments and decisions made when extensive amounts of "hard" data are available enough better than those made without such data to justify the costs involved?
4. In the last five years we've seen a great deal of progress in how we examine the results of instruction, in particular, how we determine if specific changes in knowledge have occurred. Similar developments are needed in terms of examining larger results--attaining program and institutional goals, accomplishing purpose and mission, meeting needs, solving problems. In adult education, it's the larger ends that count. Changes in knowledge are of value only if they lead to the overall development of the adult and his ability to enjoy life and manage his environment.
5. As we accept program evaluation as a generic term with many activities and many aspects of program on which it should focus, as we accept it as a toolbox full of tools rather than one tool, more work needs to be done on designing the various tools in that toolbox. First more work must be done on identifying the kinds of tools that are most needed and building some kind of a classification system that's useful. The classification system will help us organize the developments so we understand their different uses and purposes. It will help some of the confusion that occurs when we discuss the ideas of Tyler, Stufflebeam, Stake, and Scriven as though they're all dealing with the same thing. Then very serious work needs to be done on the activities that come in the various categories. For example, tools dealing with the attainment of instructional objectives are shaping up nicely. On the other hand, we're just beginning to rough out designs for evaluating needs and determining program priorities. This is a task where the programmer must take the lead rather than passing the ball to the professor.

Apparently we've moved from the era when evaluation could be compared with the first horseless carriage to the stage where it goes back to a variety of drawing boards for the kind of work that will produce 1975 models. We now know that the horseless carriage we've been using for the past 30 years can be further developed. And we have some leads about the kinds of improvements that should be made.

But improvement of the theory and the science of evaluation won't be of much value unless improvements are made in the context in which it's expected to operate. Improved theory and procedures of evaluation won't settle conflicts between administrative levels in terms of what kind of evaluation is desired. It won't solve problems caused by faulty assumptions about what it's possible for programs to do. Some of the needs related to evaluation have to be dealt with in broader contexts using research findings and good sound common sense.

Although considerable work still must be done, the progress that has been made in the past few years is extremely heartening. You'll find a good deal of help in the approaches summarized later in the monograph; ideas and help that weren't available to you 10 years ago. The wheels of progress are indeed turning.

Table 2
GROUPING OF APPROACHES TO PROGRAM EVALUATION

Group 1: Evaluation As Input Into Decision Making

- 1.1 Decision Centered Evaluation (CIPP) [Stufflebeam; Phi Delta Kappa]
- 1.2 Differential Evaluation [Tripodi, Fellin, & Epstein]
- 1.3 Discrepancy Evaluation [Provus]
- 1.4 Priority Decisions [Boyle]
- 1.5 Developmental Evaluation
 - IPI Formative Evaluation [Lindvall & Cox]
 - NewStart Evaluation System [Lamrock, Smith, & Warren]
- 1.6 Materials Evaluation
 - Trade-Off and Comparative Cost Approach [Glass]
 - Weighted Criteria Approach [Crane & Abt]
- 1.7 Participant Reaction Approaches

Group 2: Evaluation of Program Parts

- 2.1 Execution-Impact Approach [Freeman & Sherwood]
- 2.2 Actual Component Approach [Knox, Mezirow, & Darkenwald, Jr.]
- 2.3 Managerial Systems
 - Program Evaluation and Review Technique (PERT)
 - Organization As a Total System [Young]
 - Macro System Model [Alkin]
 - System Approach to Goal Setting [Van Gigh & Hill]
 - Program Contact System [Duft]
 - Management Information Systems (SEMIS)
 - Evaluating Decision Making [Hesseling]
- 2.4 Socio-Organizational Systems
 - Social Systems Models [Loomis]
 - Organizational Models [Etzioni; Schulberg & Baker]
 - Motivational Model [Lewis]

Group 3: Evaluation--Kinds of Data; Types of Activities

- 3.1 Countenance of Evaluation [Stake]
- 3.2 System Role Model [Knox]
- 3.3 Means-Ends Hierarchy [Bennett]
- 3.4 When-To-Do-It-Yourself Continuum [Alexander]

Group 4: Evaluation Processes

- 4.1 Appraisal Model [Harris]
- 4.2 Data Management [Phi Delta Kappa]
- 4.3 Natural Process Approach [Steele]
- 4.4 Monitoring Evaluation [Bruce]
- 4.5 Improvement Evaluation [Kreitlow]
- 4.6 Transactional Evaluation [Riphey]

Group 5: Results--Attainment of Objectives

5.1 Tylerian Models

The Original Tyler Model [Tyler]
 National Assessment Program
 Four-Question Approach [Gottman & Clasen]
 Adoption of the Tyler Model by Adult Education

5.2 Instructional Evaluation Approaches

Goal-Referenced Instruction [Popham & Baker]
 Evaluation As Facilitation of Learning [Bloom, Hastings, & Madaus]

5.3 Program Objectives Approaches

Criteria of Success Approach [Suchman]
 O-A-R Model [Deniston et al.]
 Criteria and Objectives [Matteson]
 Locally Directed Evaluation [Byram & Robertson]
 Ohio Model [Starr et al.]
 Program Planning Budgeting System (PPBS)
 System Variables Approaches
 NEA Model [Taba & Sawin]
 Cube Models [Hammond; Armstrong et al.]
 Information Domains [Nelson]
 Types of Evaluation [Wholey et al.]

5.4 Objectives: Categorizations and Criteria

Objectives As Systems and Parts of Systems
 Level and Time Sequence of Objectives
 Kinds of Objectives Operating Within a Program
 Characteristics of Objectives

Group 6: Results: Evaluation of Outcomes and Effects

6.1 Goal Free Evaluation [Scriven]

6.2 Zones of Results [Lamrock, Smith, & Warren]

6.3 Multiple Change Approach [Hayes, Jr.]

6.4 Multiple Dimensions of Program Effectiveness [Steele]

6.5 Effectiveness of Methods [Wilson & Gallup]

6.6 Impact Evaluation [Borus & Tash]

6.7 Public Policy [Berlak]

6.8 Institutional Evaluation [Forehand]

6.9 Social Indicators [Paulson]

6.10 Research Models [Weiss; Longest; Cain & Hollister]

6.11 Efficiency Examinations

6.12 Accountability

Section II

MATCHING EVALUATION APPROACHES TO NEEDS

As you enter your exploration of the evaluation approaches summarized in Section III, you face questions like: "How do I use evaluation approaches?" "How do I choose which one to study first?"

Using Evaluation Approaches

First, let's talk generally about use. Familiarity with various approaches to program evaluation can be useful to you in at least four ways:

1. Expanding your ideas about programming and about program evaluation so that you're better able to make evaluation an effective management tool.
2. Helping you track down sources and references that consultants and others use when talking about evaluation.
3. Providing approaches you can adapt to your own situation.
4. Providing support and ideas for building your own approaches.

Expanding Your Ideas About Programming and Program Evaluation

The help that current approaches to evaluation give you in understanding the dynamics and complexity of programming may in the long run be much more useful to you than the help that they give you in actually doing evaluation. Most of the approaches try to identify key components and the relationships among those components. The various patterns help you get a grip on the important ingredients and processes within programming. They help you identify the crucial aspects of successful programming. They help you understand the nature and role of criteria and, if the criteria are sound guides to success, result in more successful programs. A sound understanding of

the bases of program evaluation can lead to program improvement, even though evaluation isn't done in a conscious or systematic manner.

Locating Sources

Each field has its own jargon and names to be dropped. Have you ever found yourself frustrated when you attend conferences, deal with evaluation consultants, or talk with some of your colleagues who have a keen interest in conceptualizing evaluation because you aren't aware of who the names and what the ideas are? If this occurs frequently, you may want to use this monograph as a briefing device. An author index and a title index have been included at the end to help you better use it for such purposes. In all instances where the author has assigned a title, that title has been used. However, in some instances a label had to be arbitrarily assigned in the development of the monograph. Appendix A on page 220 provides a brief paragraph description of each approach organized alphabetically according to the author.

Providing You with Approaches You Can Adapt to Your Own Situation

In some instances, you'll find that an approach pretty well fits a need and you'll try to use it. In the case of the more comprehensive models, the extent to which you deal with all the parts of a particular pattern depends on the depth of the evaluation and the extent of your resources.

For example, at the everyday, "natural" level, you probably can make some observations and judgments about all of the parts of a particular pattern and the way in which they fit together. You may be able to operationalize a complete pattern without undue difficulty once the pattern is firmly in mind. However, if you're going to use a great deal of system--rigorous definition of criteria and carefully controlled gathering of data--you may

not have resources to adequately cover all of the approaches' components with the same depth of attention. This is particularly true if both the program and the evaluation approach is fairly complex. Some models such as the *countenance of evaluation*, the *system role model*, and the *means-ends hierarchy* weren't meant to be used in their entirety in normal systematic evaluation situations. They were designed to help you understand the types of data to select from in developing a meaningful evaluation.

In most instances you will select certain parts of a pattern for systematic examination. There's a growing push toward selective evaluation. For example, R. E. Brack of the University of Saskatchewan suggests that you take an eclectic approach--first identify the questions about the program that need to be answered and then select the parts of a particular model that can help deliver those answers without trying to systematically operationalize the complete model. In this situation, however, an understanding of the total pattern helps you keep the component that's receiving major attention within a total perspective of programming relationships.

Don't try to operationalize an approach by just using the material in this monograph. *Go to the original printed source.* If possible, after you've studied it thoroughly, consult the author of the model. The summaries presented here are guides to approaches rather than guides for doing evaluation.

How helpful the approaches will be when actually carried out depends on (1) the nature of your need, (2) how you conceptualize your program operations, (3) what judgments about program are most important in your situation, (4) what help in making program decisions or what other value can be gotten from a particular approach, and (5) how able you (and your evaluation consultant, if you use one) are to work from a working description rather than a precise operational blueprint.

Most of the time you will have to adapt the approach to your own situation rather than applying it "straight." However, be careful the adaptations don't distort the original model so it loses its value or becomes inappropriate.

Providing You with Support and Ideas for Building Your Own Approaches

In many instances, rather than extensively adapting a particular approach, you might be better off to construct your own, borrowing the parts of other approaches that are most useful and building patterns and processes that are appropriate to your needs.

The approach you use should stand up to tests of internal and external validity, reliability, relevance, scope, creditability, pervasiveness, and timeliness. Most of these qualities are situation specific. A model may have them in one application but not in another. If it's to be used by others, remember that you must be able to communicate it clearly to them.

Most important, you must be able to make it work so it will give reliable and believable results with a minimum of strain on the people involved.

Choosing Evaluation Approaches

Very few readers will or should try to read Section III from beginning to end at one sitting. (If you should want to examine all of the approaches, you'll probably function better by working through a group at a time, starting with a particular need or interest in evaluation and going directly to the approach or approaches that best relate to that need.)

Caution: Don't be channeled completely by one need or view of what evaluation is. Remember that evaluation is a generic phenomenon and that it includes several approaches. *Don't search for the one way to do evaluation. Do search for the range of approaches that will best address your varied needs in program evaluation.*

Start your evaluation approaches by outlining the kinds of needs that you and your unit have that evaluation can help with. It may be helpful for you to take a dual track. On one sheet, list the *kinds of evaluation frameworks that you need if your kit of evaluative tools is to be complete*. On another sheet, list the *specific problems you and your unit face where evaluation can be of help*.

Types of Evaluation Frameworks

A program unit needs a group of frameworks to be applied frequently and consistently--for example, a framework for doing the kinds of evaluation that are essential throughout the program process and for dealing with everyday management questions; a framework that views the program through a participant's eyes; a framework that monitors the effectiveness of the various components of a program. These types of frameworks don't substitute for each other. They need to be tied together through a common philosophy of what evaluation is and an understanding of the general process of evaluation and subprocesses such as appraisal and data management. In addition, the program unit may also need some frameworks for dealing with more specific tasks like: choosing visual aids, texts, and other support materials; establishing priorities for programs; introducing change into a system. Table 3 provides examples of how various approaches summarized in the monograph relate to needs for general kinds of evaluative frameworks.

Some frameworks are more important to one program than to another or are more important at a particular point in time. For example, programs for disadvantaged adults more than other programs may need frameworks for seeing how program participants evaluate the program or may need to use such frameworks extensively during the first two years of operation but give less attention to them in later years. Programs that develop all of their own materials won't need any frameworks for evaluating materials to be purchased.

Table 3
COMMONLY NEEDED EVALUATION FRAMEWORKS

Most programmers and program units need more than one evaluation framework in their tool kit. Different approaches accomplish different things.

<u>Types of evaluation frameworks needed by program units</u>	<u>Examples of approaches included in this monograph</u>	<u>On page:</u>
<i>Frameworks that help make major decisions and guides overall program management</i>	Decision Centered Evaluation (CIPP)	78
	Differential Evaluation	81
	Criteria of Success Approach	171
	O-A-R Model	173
	Program Planning Budgeting System (PPBS)	178
<i>Frameworks for examining the impact and larger results of programs</i>	Multiple Change Approach	198
	Multiple Dimensions of Program Effectiveness	200
	Impact Evaluation	203
	Social Indicators	210
	Means-Ends Hierarchy	129
	Public Policy	205
	Research Models	212
	Criteria of Success Approach	171
	Effectiveness of Methods	202
	Zones of Results	195
	Efficiency Examinations	215
	Accountability	217
	Ohio Model	177
	Criteria and Objectives	174
	Types of Evaluation	181
<i>Frameworks that guide the organization and use of program components</i>	Actual Component Approach	104
	Program Evaluation and Review Technique (PERT)	107
	Macro System Model	110
	Organization As a Total System	109
	Countenance of Evaluation	120
	System Role Model	124
	Social Systems Models	116
	Organizational Models	117
	Program Contact System	112
	Management Information Systems (SEMIS)	114
	System Variables Approaches	179

<u>Types of evaluation frameworks needed by program units</u>	<u>Examples of approaches included in this monograph</u>	<u>On page:</u>
<i>Frameworks for viewing the program through the eyes of the participant</i>	Participant Reaction Approaches Locally Directed Evaluation Motivational Model Transactional Evaluation Institutional Evaluation	98 176 118 150 207
<i>Frameworks that guide evaluative processes</i>	Appraisal Model Data Management Natural Process Approach Monitoring Evaluation Locally Directed Evaluation When-To-Do-It-Yourself Continuum	135 138 143 146 176 131
<i>Frameworks for examining the results of instruction</i>	Tylerian Models Goal-Referenced Instruction Evaluation As Facilitation of Learning Goal Free Evaluation Institutional Evaluation Objectives: Categorizations and Criteria Adoption of the Tyler Model by Adult Education	154 164 168 192 207 184 157
<i>Frameworks that serve specific purposes</i>	Discrepancy Evaluation Priority Decisions Developmental Evaluation Materials Evaluation	84 86 89 95

However, adult basic education which draws on many commercially prepared materials may find such a framework essential to effective operation. Some programs may need frameworks for examining instructional results more than they need frameworks for examining larger results. On the other hand, other programs for disadvantaged adults need to concentrate on exploring major impact and value, and if resources don't stretch to cover both, give less attention to the specific results of instruction.

Each program unit has to decide which frameworks are most relevant and important to the operation of the particular unit, and as a result, which frameworks should be extensively operationalized and which should be drawn on periodically or at time of greatest need. Even though a program unit restricts itself to two or three general frameworks, if this restriction is made after a full survey of the kinds of program evaluation that could and perhaps should be operating, there is a much better understanding of what the particular framework can (and can't) be expected to do.

Use of an Approach to Deal with a Specific Problem

A second approach to choosing evaluation approaches is that of using an approach to deal with a specific problem. For example, the administrator recognizes that changes are needed in a program that staff are apt to resist. In such an instance, an approach like *transactional evaluation* might be particularly helpful. Or, if staff feel uneasy about a particular program phase but are unable to really diagnose the problem, *discrepancy evaluation* could be helpful. Or, if staff need a challenge plus a pat on the back to give them further impetus, the *improvement evaluation* approach would be useful. In most instances, the problem probably will result in one of the major frameworks accepted by the program unit being applied, but in other

instances like the ones given above, an additional evaluative approach not commonly part of the tool kit may be selected.

Table 4 includes examples of problem situations that sometimes arise and indicates the approaches within the monograph that might be useful in dealing with those situations. They have been loosely grouped into three types: problems in programming, problems in program management, and problems in evaluation.

Table 4

MATCHING EVALUATION APPROACHES TO PROBLEMS AND NEEDS

If you're having problems in *PROGRAMMING* . . .

<u>With:</u>	<u>Look for ideas in:</u>	<u>On page:¹</u>
<i>Setting priorities</i>	Priority Decisions	86
	Decision Centered Evaluation (CIPP) [context]	78
<i>Choosing among program possibilities</i>	Decision Centered Evaluation (CIPP) [context]	78
	Priority Decisions	86
<i>Determining the kind of results you're aiming for as you design the program</i>	Means-Ends Hierarchy	129
	System Role Model	124
	Multiple Change Approach	198
	Multiple Dimensions of Program Effectiveness	200
	Impact Evaluation	203
	Research Models	212
	Social Indicators	210
	Goal-Referenced Instruction	164
	Evaluation As Facilitation of Learning	168
	Objectives: Categorizations and Criteria	184
<i>Developing objectives</i>	System Approach to Goal Setting	111
	Decision Centered Evaluation (CIPP) [context]	78
	Priority Decisions	86
	Objectives: Categorizations and Criteria	184
	Tylerian Models	154
<i>Determining the right level of objectives</i>	Objectives: Categorizations and Criteria	184
	Criteria of Success Approach	171

¹This table is divided into three main sections--problems in programming, problems in program management, and problems in doing evaluation. Groupings are general and not mutually exclusive. Models suggested as sources of ideas for a particular task are examples and not the only relevant ones.

If you're having problems in *PROGRAMMING* . . . (cont.)

<u>With:</u>	<u>Look for ideas in:</u>	<u>On page:</u>
<i>Selecting content and focus</i>	Decision Centered Evaluation (CIPP) [context]	78
	Participant Reaction Approaches	98
	Program Contact System	112
<i>Identifying what participants see as appropriate program activities</i>	Participant Reaction Approaches	98
<i>Examining factors affecting participation and learning</i>	Decision Centered Evaluation (CIPP) [context]	78
	Social Systems Models	116
<i>Developing a project or program plan</i>	IPI Formative Evaluation	89
	NewStart Evaluation System	91
	Program Evaluation and Review Technique (PERT)	107
<i>Developing a pilot project</i>	Discrepancy Evaluation	84
	NewStart Evaluation System	91
	Social Systems Models	116
	Tylerian Models	154
<i>Determining how you can use scarce resources most effectively in the program</i>	Decision Centered Evaluation (CIPP) [input]	78
	Differential Evaluation [efficiency]	81
<i>Making sure that all of the elements of the program are appropriate</i>	Appraisal Model	135
<i>Understanding how various aspects of the program fit together</i>	Actual Component Approach	104
	Program Evaluation and Review Technique (PERT)	107

If you're having problems in *PROGRAMMING* . . . (cont.)

<u>With:</u>	<u>Look for ideas in:</u>	<u>On page:</u>
<i>Working effectively with others in the program</i>	Motivational Model Transactional Evaluation Organizational Models	118 150 117
<i>Choosing among textbooks, audio-visual packages, etc.</i>	Trade-Off and Comparative Cost Approach Weighted Criteria Approach	95 96
<i>Preparing teaching plans</i>	Goal-Referenced Instruction Evaluation As a Facilitation of Learning Objectives: Categorizations and Criteria	164 168 184
<i>Determining whether a pro- gram is progressing toward its goals</i>	Monitoring Evaluation	146
<i>Finding out why a program isn't producing as much as you expect it to</i>	Discrepancy Evaluation	84
<i>Finding out how well your first programming efforts are achieving</i>	Differential Evaluation [initiation and contact] Monitoring Evaluation	81 146
<i>Clarifying what kind of results to expect from a new program</i>	Goal Free Evaluation Research Models [search evaluation]	192 212
<i>Understanding how program components interact to produce results</i>	Actual Component Model Macro System Model Countenance of Evaluation System Role Model Program Evaluation and Review Technique (PERT) System Variables Approaches	104 110 120 124 107 179

If you're having problems in *PROGRAMMING* . . . (cont.)

<u>With:</u>	<u>Look for ideas in:</u>	<u>On page:</u>
<i>Testing a particular instructional sequence to see what it's achieving</i>	Tylerian Models Goal-Referenced Instruction Evaluation As a Facilitation of Learning	154 164 168
<i>Selecting the most effective method for a particular task</i>	Effectiveness of Methods	202
<i>Determining the effect of a program on a community</i>	Impact Evaluation Social Indicators Trade-Off and Comparative Cost Approach	203 210 95
<i>Identifying whether you're reaching the right clientele</i>	Execution-Impact Approach Participant Reaction Approaches Adoption of the Tyler Model by Adult Education	102 98 157
<i>Revising a pilot program and developing a guide for others to follow</i>	Execution-Impact Approach Discrepancy Evaluation Macro System Model NewStart Evaluation System System Role Model	102 84 110 91 124
<i>Getting other people's ideas on how your program can be improved</i>	Appraisal Model Improvement Evaluation Participant Reaction Approaches	135 148 98
<i>Improving on-the-spot judgments as a program progresses</i>	Natural Process Evaluation	143
<i>Examining the results of programs</i>	Decision Centered Evaluation (CIPP) [context] Trade-Off and Comparative Cost Approach Execution-Impact Approach Countenance of Evaluation System Role Model Means-Ends Hierarchy Tylerian Models Instructional Evaluation	78 95 102 120 124 129 154 207

If you're having problems in *PROGRAMMING . . .* (cont.)

<u>With:</u>	<u>Look for ideas in:</u>	<u>On page:</u>
<i>Examining the results of programs (cont.)</i>	Program Objectives Approaches	171
	Goal Free Evaluation	192
	Zones of Results	195
	Multiple Change Approach	198
	Multiple Dimensions of Program Effectiveness	200
	Impact Evaluation	203
	Institutional Evaluation	207
	Social Indicators	210
	Research Models	212
	Efficiency Examinations	215
<i>Identifying key elements contributing to the success or failure of a program</i>	Actual Component Approach	104
	Macro System Model	110
	Improvement Evaluation	148
	System Variables Approaches	179
	Countenance of Evaluation	120
	Means-Ends Hierarchy	129
<i>Preparing a program report</i>	Data Management [communication]	138
	Trade-Off and Comparative Cost Approach	95
<i>Deciding what to emphasize in a report prepared for administrators and influentials</i>	Multiple Change Approach	198
	Multiple Dimensions of Program Effectiveness	200
	Impact Evaluation	203
<i>If you're having problems in PROGRAM MANAGEMENT . . .</i>		
<i>Guiding a program that involves several staff members with roles dependent on prior work by others</i>	Program Evaluation and Review Technique (PERT)	107
	Organizational Models	117
<i>Understanding the kind of decisions involved in programming</i>	Decision Centered Evaluation (CIPP) [input]	78
	NewStart Evaluation System	91
	Evaluating Decision Making	115
<i>Getting the most out of limited resources</i>	Decision Centered Evaluation (CIPP) [input]	78
	IPI Formative Evaluation	89
	Participant Reaction Approaches	98
	Trade-Off and Comparative Cost Approach	95

If you're having problems in *PROGRAM MANAGEMENT* . . . (cont.)

<u>With:</u>	<u>Look for ideas in:</u>	<u>On page:</u>
<i>Deciding whether ideas for programs are good</i>	Priority Decisions	86
	Decision Centered Evaluation (CIPP) [input]	78
	IPI Formative Evaluation	89
	NewStart Evaluation System	91
	Goal Free Evaluation	192
	Trade-Off and Comparative Cost Approach	95
<i>Supervising activities of several para professionals</i>	Execution-Impact Approach	102
	System Role Model	124
	Motivational Model	118
<i>Securing additional funding</i>	Data Management [Communication]	138
	Execution-Impact Approach	102
	Differential Evaluation [effectiveness, efficiency]	81
	Program Objectives Approaches	171
	Multiple Change Approach	198
	Multiple Dimensions of Program Effectiveness	200
	Accountability	217
	Trade-Off and Comparative Cost Approach	95
	Decision Centered Evaluation (CIPP)	78
<i>Developing long-range goals and mission</i>	System Approach to Goal Setting	111
	Priority Decisions	86
	Multiple Change Approach	198
	Objectives: Categorizations and Criteria	184
	IPI Formative Evaluation	89
<i>Reviewing long-range goals</i>	IPI Formative Evaluation	89
	Multiple Change Approach	198
	Priority Decisions	86
	Goal Free Evaluation	192
<i>Improving teamwork within your unit</i>	Social Systems Models	116
	Motivational Model	118
	Organization As a Total System	109
	Transactional Evaluation	150

If you're having problems in *PROGRAM MANAGEMENT* . . . (cont.)

<u>With:</u>	<u>Look for ideas in:</u>	<u>On page:</u>
<i>Understanding the stages of programming</i>	Differential Evaluation	81
	Program Contact System	112
<i>Understanding relationship of evaluation to decision making</i>	Decision Centered Evaluation (CIPP)	78
	NewStart Evaluation System	91
	Trade-Off and Comparative Cost Approach	95
<i>Defining program efficiency</i>	Efficiency Examinations	215
	Differential Evaluation	81
	Criteria of Success Approach	171
	Execution-Impact Approach	102
<i>Settling disputes when two or more methods or plans are advocated for doing the same thing</i>	Decision Centered Evaluation (CIPP) [input]	78
	Differential Evaluation	81
	Effectiveness of Methods	202
<i>Designing a management information system for your unit</i>	Management Information Systems (SEMIS)	114
	System Role Model	124
	Data Management	138
<i>Designing accountability strategies</i>	Accountability	217
	Decision Centered Evaluation (CIPP)	78
	Locally Directed Evaluation	176
	Types of Evaluation	181
	Management Information Systems (SEMIS)	114
	Data Management	138
<i>Developing an organization and climate that will improve programming</i>	Motivational Model	118
	Organizational Models	117
	Social Systems Models	116
	Transactional Evaluation	150
<i>Identifying and weighing groups and other pressures that must be considered in programming</i>	Priority Decisions	86
	Social Systems Models	116

If you're having problems in *PROGRAM MANAGEMENT* . . . (cont.)

<u>With:</u>	<u>Look for ideas in:</u>	<u>On page:</u>
<i>Getting new ideas adopted by the staff</i>	Motivational Model	118
	Discrepancy Evaluation	84
	Transactional Evaluation	150
<i>Organizing support resources effectively</i>	Organizational Models	117
	Program Evaluation and Review Technique (PERT)	107
<i>Increasing operational efficiency</i>	Efficiency Examinations	215
	Differential Evaluation	81
	Program Evaluation and Review Technique (PERT)	107
	Program Planning Budgeting System (PPBS)	178
	Criteria of Success Approach	171
	Trade-Off and Comparative Cost Approach	95
<i>Cooperating with content experts</i>	Program Contact System	112
	Motivational Model	118
	Transactional Evaluation	150
	Discrepancy Evaluation	84
<i>Using a systems approach to organizing programming</i>	Organization As a Total System	109
	Actual Component Approach	104
	Macro System Model	110
	Countenance of Evaluation	120
	System Role Model	124
	Program Evaluation and Review Technique (PERT)	107
	System Variables Approaches	179
<i>Developing criteria to guide the work of your unit</i>	Criteria of Success Approach	171
	Natural Process Approach	143
	Differential Evaluation	81
	Discrepancy Evaluation	84
	Multiple Dimensions of Program Effectiveness	200
	Public Policy	205
	Institutional Evaluation	207
	Trade-Off and Comparative Cost Approach	95

If you're having problems in *PROGRAM MANAGEMENT* . . . (cont.)

<u>With:</u>	<u>Look for ideas in:</u>	<u>On page:</u>
<i>Considering input, process, and output relationships of various parts of your program</i>	O-A-R Model	173
	Actual Component Approach	104
	Organization As a Total System	109
	Macro System Model	110
	System Approach to Goal Setting	111
	Countenance of Evaluation	120
	System Role Model	124
<i>Developing a multicourse curricula or a multiactivity program</i>	Institutional Evaluation	207
	Decision Centered Evaluation (CIPP)	78
	Trade-Off and Comparative Cost Approach	95
<i>Increasing the contributions of volunteers</i>	Motivational Model	118
	Social Systems Models	116
<i>Improving everyday operations</i>	Natural Process Approach	143
	Trade-Off and Comparative Cost Approach	95
<i>Determining whether programming activities are on target in terms of the unit's mission</i>	Monitoring Evaluation	146
<i>Using information in making program decisions</i>	Data Management	138
	Natural Process Approach	143
<i>Assigning resources to staff units</i>	Program Planning Budgeting System (PPBS)	178
	Decision Centered Evaluation (CIPP)	78
	Efficiency Examinations	215
	Accountability	217
	Multiple Dimensions of Program Effectiveness	200
	Program Objectives Approaches	171
	Trade-Off and Comparative Cost Approach	95

If you're having problems in *PROGRAM MANAGEMENT* . . . (cont.)

<u>With:</u>	<u>Look for ideas in:</u>	<u>On page:</u>
<i>Managing the teaching-learning transaction</i>	Evaluation As a Facilitation of Learning	168
	Zones of Results	195
	System Variables Approaches	179
	Objectives: Categorizations and Criteria	184
	Monitoring Evaluation	146
	Natural Process Approach	143
	Trade-Off and Comparative Cost Approach	95
<i>Developing a personnel evaluation system</i>	Multiple Dimensions of Program Effectiveness	200
	Appraisal Model	135
	Natural Process Approach	143
	Weighted Criteria Approach	96
	Trade-Off and Comparative Cost Approach	95
<i>Analyzing weaknesses and problems in the operation of the unit</i>	Discrepancy Evaluation	84
	Program Evaluation and Review Technique (PERT)	107
	Decision Centered Evaluation (CIPP)	78
	Appraisal Model	135
	Improvement Evaluation	148
	Trade-Off and Comparative Cost Approach	95
<i>Working with an advisory committee</i>	Locally Directed Evaluation	176
	Decision Centered Evaluation (CIPP)	78
	System Approach to Goal Setting	111
	Social Systems Models	116
	Accountability	217
	Data Management [communication]	138
	Monitoring Evaluation	146
	Natural Process Approach	143
	Appraisal Model	135
<i>Improving use of time and other resources</i>	Program Evaluation and Review Technique (PERT)	107
	Decision Centered Evaluation (CIPP)	78
	Trade-Off and Comparative Cost Approach	95

If you're having problems in *PROGRAM MANAGEMENT* . . . (cont.)

<u>With:</u>	<u>Look for ideas in:</u>	<u>On page:</u>
<i>Developing budgets</i>	Program Planning Budgeting System (PPBS)	178
	Impact Evaluation	203
	Decision Centered Evaluation (CIPP)	78
	Differential Evaluation	81
<i>Guiding and controlling the total programming process</i>	Differential Evaluation	81
	Organization As a Total System	109
	System Role Model	124
	Program Contact System	112
<i>Helping staff use evaluation as a management tool</i>	Natural Process Approach	143
	Monitoring Evaluation	146
	Decision Centered Evaluation (CIPP)	78
	Differential Evaluation	81
	Discrepancy Evaluation	84
	Trade-Off and Comparative Cost Approach	95

If you're having problems in *EVALUATION* . . .

<i>Improving staff attitude toward evaluation</i>	Natural Process Approach	143
	Discrepancy Evaluation	84
	Decision Centered Evaluation (CIPP)	78
	Motivational Model	118
	Transactional Evaluation	150
<i>Distinguishing between describing and evaluating program results</i>	Countenance of Evaluation	120
<i>Determining results when objectives were poorly stated or changed during the program</i>	Goal Free Evaluation	192
	Multiple Dimensions of Program Effectiveness	200
<i>Developing criteria for judging the program</i>	Appraisal Model	135
	Natural Process Approach	143
	Discrepancy Evaluation	84
	Differential Evaluation	81
	Criteria of Success Approach	171
	Execution-Impact Approach	102
	Multiple Dimensions of Program Effectiveness	200

If you're having problems in *EVALUATION* . . . (cont.)

<u>With:</u>	<u>Look for ideas in:</u>	<u>On page:</u>
<i>Developing criteria for judging the program (cont.)</i>	Institutional Evaluation Efficiency Examinations Accountability	207 215 217
<i>Dealing with several criteria at one time</i>	Weighted Criteria Approach	96
<i>Forming judgments about programs</i>	Natural Process Approach Appraisal Model Decision Centered Evaluation (CIPP) Trade-Off and Comparative Cost Approach	143 135 78 95
<i>Involving lay people in evaluation</i>	Locally Directed Evaluation Participant Reaction Approaches Discrepancy Evaluation	176 98 84
<i>Organizing a comprehensive program review, accreditation team, etc.</i>	Appraisal Model Improvement Evaluation Actual Component Approach System Role Model Organization As a Total System Decision Centered Evaluation (CIPP) Multiple Dimensions of Program Effectiveness Goal Free Evaluation	135 148 104 124 109 78 200 192
<i>Preparing for an accreditation team, monitoring team, or comprehensive review</i>	Appraisal Model Execution-Impact Approach Multiple Dimensions of Program Effectiveness Natural Process Approach Data Management [communication]	135 102 200 143 138
<i>Determining what parts of the program to include in your evaluation</i>	Decision Centered Evaluation (CIPP) Actual Component Approach Appraisal Model Improvement Evaluation Countenance of Evaluation System Role Model Means-Ends Hierarchy Natural Process Approach Multiple Dimensions of Program Effectiveness	78 104 135 148 120 124 129 143 200

If you're having problems in *EVALUATION* . . . (cont.)

<u>With:</u>	<u>Look for ideas in:</u>	<u>On page:</u>
<i>Determining what parts of the program to include in your evaluation (cont.)</i>	Multiple Change Approach System Variables Approaches O-A-R Model	198 179 173
<i>Identifying appropriate data bases and sources of data</i>	NewStart Evaluation System Information Domains	91 181
<i>Selecting techniques to carry out specific evaluation tasks</i>	Differential Evaluation	81
<i>Categorizing kinds of judgments that need to be made about a program</i>	Differential Evaluation Natural Process Approach Criteria of Success Approach Criteria and Objectives Execution-Impact Approach	81 143 171 174 102
<i>Deciding whether to consult an evaluation expert</i>	When-To-Do-It-Yourself Continuum	131
<i>Determining whether an extensive research effort is appropriate</i>	Types of Evaluation Decision Centered Evaluation (CIPP)	181 78
<i>Handling a good deal of unrelated data and making sense out of how it fits into patterns</i>	Means-Ends Hierarchy System Role Model Countenance of Evaluation Actual Component Approach	129 124 120 104
<i>Setting performance standards</i>	Countenance of Evaluation Appraisal Model Natural Process Approach Discrepancy Evaluation Goal-Referenced Instruction Multiple Dimensions of Program Effectiveness Criteria of Success Approach Differential Evaluation	120 135 143 84 164 200 171 81
<i>Determining the kind of data needed</i>	Natural Process Approach Countenance of Evaluation Means-Ends Hierarchy Decision Centered Evaluation (CIPP) Research Models	143 120 129 78 212

If you're having problems in *EVALUATION* . . . (cont.)

<u>With:</u>	<u>Look for ideas in:</u>	<u>On page:</u>
<i>Establishing or evaluating sampling procedures</i>	Data Management Research Models	138 212
<i>Establishing or evaluating treatment procedures</i>	Data Management Research Models	138 212
<i>Establishing or evaluating measurement procedures</i>	Data Management Research Models	138 212
<i>Examining assumptions related to data</i>	Data Management	138
<i>Identifying flaws and limitations in data as a basis for arguing against an unfavorable evaluation</i>	Data Management	138
<i>Identifying flaws and limitations in data as a basis for assessing the probable accuracy of the data</i>	Data Management	138
<i>Developing evaluation reports</i>	Data Management [communication] System Role Model Trade-Off and Comparative Cost Approach	138 124 95
<i>Getting evaluation findings used</i>	Data Management [communication] System Role Model Transactional Evaluation Discrepancy Evaluation	138 124 150 84
<i>Evaluating evaluation</i>	Data Management Trade-Off and Comparative Cost Approach	138 95

Section III

SUMMARIES OF CONTEMPORARY PROGRAM EVALUATION APPROACHES

This section provides a brief summary of contemporary approaches to program evaluation. The approaches are grouped into the following categories according to the primary focus of the approach:

- Group 1. Evaluation as input into decision making.
- Group 2. Evaluation of program parts.
- Group 3. Evaluation--kinds of data; types of activities.
- Group 4. Evaluation processes.
- Group 5. Results--attainment of objectives.
- Group 6. Results--evaluation of outcomes and effects.

Groups 5 and 6 emphasize examining the results of programs. The other four groups focus on results and other elements of the program system in such a way that the information can be used in developing and improving programs while they're in progress. A very brief introduction is provided to each group of approaches.

The following information is given for each approach included within a group: field of origin, the essence of the approach, general suggestion as to use, place to start when using it, description of the approach, and a printed source where a more complete presentation can be found.

The summaries should be treated as a "reference to" rather than a "reference for." They'll help you decide which ones you'll want to get and study in greater depth or which authors you'll want to contact directly. The summaries don't provide a complete reference on how to use the particular approach.

GROUP 1

EVALUATION AS INPUT INTO DECISION MAKING

OVERVIEW:

What kinds of decisions do you have to make about programs? What kinds of decisions do you make while you're programming? This group of models emphasizes evaluation's role in decision making. The models: (1) pattern types of decisions that have to be made, (2) identify program sequences that should be evaluated, (3) outline evaluation questions that need to be answered, or (4) identify criteria that should be applied.

EXAMPLES INCLUDED:

- The *decision centered evaluation* model identifies four types of evaluation that correspond to the four major kinds of program decisions.
- *Differential evaluation* identifies three types of criteria that need to be applied to each of three programming stages.
- *Discrepancy evaluation* outlines a process by which standards are set for program activities and results and gaps identified by comparing actual performance with those standards.
- The *priority decisions* approach describes six categories of criteria which may be important in deciding among needs and program alternatives.
- The two examples of *developmental evaluation* describe the process used in developing program prototypes that will be used many times in many situations.
- The two examples of models for *materials evaluation* provide guides for choosing among packaged curricula.
- *Assessing participant reactions* provides a summary of efforts of this type.

GENERAL ADVANTAGES:

- Models like these help evaluation work for you in improving programs.
- They provide ways of dealing with decisions.
- They channel evaluation's energy back into ongoing programming.

GENERAL CAUTIONS:

- The models provide examples of patterns. Your view of programming may not coincide exactly. For example, you may want to add other program stages, major decision categories, or kinds of criteria.
- As much emphasis should be placed on using as is placed on securing information.

1.1 DECISION CENTERED EVALUATION (CIPP)

ORIGIN	Metropolitan and other local school systems.
ESSENCE	Four types of evaluation-- <i>context, input, process, and product</i> --correspond to four major kinds of decisions-- <i>planning, structuring, implementing, and recycling</i> .
USE	Used throughout programming, but particularly useful in: • Viewing total program over a period of time. • Developing new programs. • Understanding the decision dynamics within programs.
STARTING POINT	Set decision parameters. What decisions must be made, by whom, at what time? What information is needed for judging decision alternatives?
DESCRIPTION	Educational evaluation is defined as the process of delineating, obtaining, and providing information useful for <i>judging decision alternatives</i>. It can be helpful at all four stages of decision making--<i>awareness, design, choice, and action</i>. It facilitates decision making by establishing criteria for distinguishing among alternatives. The CIPP model identifies four major types of evaluation that correspond to the four types of decisions involved in programming. C = <i>Context</i>--serves planning decision. I = <i>Input</i>--serves decisions about designs and resources. P = <i>Process</i>--serves decisions that control operations. P = <i>Product</i>--serves decisions about results and recycling. Context Evaluation. Provides the rationale for determining objectives and setting priorities. It defines the relevant environment, describes the desired and actual conditions pertaining to that environment, and identifies unmet needs and unused opportunities. It considers such things as behavior of students, curriculum, staff strengths and weaknesses, facilities, financing, and the community. It examines present and emerging value systems and provides means for setting priorities. It examines the amount of change needed.

Input Evaluation. Provides information for determining how to use resources to meet program goals. It deals with relevance, practicability, costs, projected effectiveness, and superiority alternatives. It involves identifying and assessing: (1) relevant capabilities of the system, (2) strategies for achieving program goals, and (3) designs for implementing a selected strategy. Two or more ways of carrying out the same program are analyzed in terms of expected costs and benefits. Has this kind of plan worked in the past? What are its underlying assumptions and can they be met? What side effects may be produced? How do people react to the approach? Can the plan actually be carried out successfully?

Process Evaluation. Provides feedback. It's concerned with the extent of operational efficiency including effects on rest of the system and feasibility. It detects or predicts defects in the procedural design and its implementation, provides information for decisions that are part of implementation, and maintains a record of the procedures followed.

It must be continually alert in advance to the kinds of decisions the teacher will be making. It's concerned with such things as interpersonal relationships among staff and students, communication channels, logistics, resources, time schedules as potential causes of failures. The record of procedures helps to later analyze why certain results occurred.

Product Evaluation. Measures and interprets attainments during and at the end of the program. It's concerned with main effects, side effects, costs, and superiority. Product evaluation involves establishing criteria, taking measurements, comparing measurement results with absolute or relative standards, and making rational interpretations of the outcomes using context, input, and process information.

Criteria may be either *consequential* or *instrumental*. Consequential criteria pertain to the fundamental long-range conditions being sought. Instrumental criteria deal with accomplishments at an intermediate level that contribute to the ultimate attainment.

The four types can be used independently or in combinations. They can be used to facilitate decisions as the program progresses (formative evaluation) or in retrospective analysis of the quality of decisions that were made and implemented (summative evaluation and accountability).

Formal, structured, comprehensive evaluation is most apt to be needed when the decisions involve major changes in areas where decision makers have little experience and little information.

Evaluation models should be grounded in decision making and change theory. Evaluation designs should satisfy criteria of scientific adequacy (internal and external validity, reliability, and objectivity), of practical use (relevance, importance, scope, creditability, timeliness, and pervasiveness), and of prudential worth (efficiency).

SOURCES

Phi Delta Kappa National Study Committee on Evaluation. *Educational Evaluation and Decision Making*. Itasca, Illinois: F. E. Peacock Publishers, Inc., 1971.

Stufflebeam, Daniel L. "Evaluation As Enlightenment for Decision-Making." Address at working conference on assessment theory, the Commission of Assessment of Educational Outcomes, the Association for Supervision and Curriculum Development, Sarasota, Florida, January, 1968.

Stufflebeam, Daniel L. "The Relevance of the CIPP Evaluation Model for Educational Accountability." *Journal of Research and Development in Education*, V (February, 1971), 19-25.

1.2 DIFFERENTIAL EVALUATION

ORIGIN	Social fields.
ESSENCE	Systematic accumulation of facts about program achievement in terms of <i>effort</i> , <i>effectiveness</i> , and <i>efficiency</i> within any of the stages of development-- <i>initiation</i> , <i>contact</i> , and <i>implementation</i> .
USE	A means by which administrators and program personnel can decide where evaluation investments will be apt to give the greatest payoff in program improvement. Particularly useful in programs that haven't found the right routine or that aren't accomplishing as much as they would like.
STARTING POINT	Determine key people. Secure understanding and commitment. Examine possible consequences. Decide on the kinds of evaluative questions that will provide the most valuable information.
DESCRIPTION	Evaluation is a management technique for the systematic feedback of information to be used to improve social programs. Programmers who are committed to improving societal conditions through programs must contend with pressures from a variety of sources. In many instances, they must make quick decisions in ambiguous situations. The right information at the right time can be helpful. As the main consumer of evaluation studies, administrators should play a major role in planning evaluation. The process of evaluation involves getting systematic information within a context of social relationships. The people involved may hold different ideologies and value systems and various degrees of vested interests. A crucial element in obtaining useful evaluations is the extent to which key people can accommodate one another. Examining political realities is as important in evaluation as it is in planning. Planning for social program evaluation depends a great deal on specifying program requisites within the different stages of development. <i>Differential evaluation</i> is a process of asking different evaluation questions for each stage of program development and then choosing those evaluation techniques most appropriate to the evaluation objectives.

The three categories that define the kinds of questions are:

Program Efforts. Amounts and kinds of program activities

necessary with a particular stage of development; amount of resources invested and how those resources were applied. Evaluation is concerned with the type and quantity of program activities.

Program Effectiveness. Extent to which goals have been achieved; analysis of both desirable and undesirable unanticipated consequences. Evaluation is concerned with whether these goals have been attained as a result of program efforts.

Program Efficiency. The relationship between efforts and effectiveness. It's the ratio of effectiveness to effort and is concerned with the relative costs of achieving results. Costs include manpower, time, money, and facilities. Evaluation is concerned with the relative costs of achieving outcomes.

The three program development stages are:

Program Initiation. Stage where ideas are translated into a plan of action and necessary resources are secured.

Program Contact. Stage where contact is made with target clientele; includes delivery system and analysis of obstacles, and establishing relationships with other programs in the community.

Program Implementation. Energy is applied toward the goals of the program. Concern is with results. Criteria for follow-up activities and possible program termination are specified.

In these three stages:

1. Longevity and complexity aren't indicative of a stage of program development.
2. Different programs within the same agency may be in different stages of development.
3. Progression through the stages isn't always distinct and linear.
4. Some programs may be concerned entirely with contact.

The categories and stages produce questions like the following:

Program Initiation Stage. To what extent are efforts made to locate anticipated target populations (effort)? To what extent does the program meet a community need (effectiveness)? To what extent do staff members have competing objectives that are at cross purposes with one another (efficiency)?

Program Contact Stage. What amounts of time and energy are devoted to finding resources that could increase the number of program contacts (effort)? To what extent is the intended target population represented in those who are designated as program beneficiaries (effectiveness)? What are the relative costs of using different means for contacting clients (efficiency)?

Program Implementation Stage. How much effort is devoted to the specification of criteria for program termination and necessary follow-up activities (effort)? What's the relative effectiveness of the program compared with other programs that have similar objectives (effectiveness)? What's the relation of costs of program efforts to the benefits achieved (efficiency)?

The questions determine the kind of information-gathering technique to use. The most common information-gathering techniques can be categorized as follows:

Monitoring Techniques. Procedures that are used for direct review of program operations--accountability audit, time and motion studies, administrative audit.

Social Research Techniques. Procedures used for developing, modifying, and expanding knowledge about the program--experiment, survey, case study.

Cost Analytic Techniques. Procedures used to appraise the relative value of program benefits in relation to program costs--cost accounting, cost-benefit analysis, cost-outcome analyses, and operations research.

SOURCE

Tripodi, Tony, Phillip Fellin, and Irwin Epstein. *Social Program Evaluation*. Itasca, Illinois: F. E. Peacock Publishers, Inc., 1971.

1.3 DISCREPANCY EVALUATION

ORIGIN	Pittsburgh public school system and other education projects.
ESSENCE	Program personnel set <i>standards</i> for activities at each programming stage-- <i>design, installation, process, and product</i> --and for the results expected. Actual performance is compared with the standards and <i>discrepancies</i> or areas for improvement identified.
USE	Particularly helpful in: ● Checking out pilot or other new programs. ● Diagnosing programs that aren't functioning well. ● Improving program efficiency. ● Helping new personnel better understand program development.
STARTING POINT	Determine the key stages or points where standards should be applied. Devise the means by which the standards will be set.
DESCRIPTION	The following rationale underlies the approach to setting <i>standards</i>. Self-evaluation is a powerful device, perhaps the essential mechanism in changing individual and group behavior. A rational person will change his standards if they don't serve a value or are counterproductive to a higher value. An autonomous person must want to change his standards before he will change his behavior. Standards can be changed through the analysis and synthesis of phenomena, insight, and the reorganization of a perceptual field, such as through the problem-solving process. It's possible to get a person or group to change standards by showing discrepancies between performance and the assumptions and values that gave rise to the standard. A non-programmer (evaluation staff member) plays an important role in catalyzing the process by raising questions, supplying information that the programmer doesn't have time to collect and examine, and helping the programmer clarify values and biases. A program is a dynamic phenomenon that's made up of stages--<i>design, installation, process, and product</i>. Each stage involves <i>inputs, processes, and outputs</i>. At each stage, evaluation includes: (1) defining program standards,

(2) determining whether a discrepancy exists between some aspect of program performance and the standards governing that aspect of the program, and (3) using discrepancy information either to change performance or to change program standards.

When the *discrepancies* have been identified, the next sequence involves determining: (1) why there's a discrepancy, (2) what corrective actions are possible, and (3) what corrective action is best. Three elements are necessary in this process: (1) criteria for identifying relevant information and knowing how to interpret it based on some standard for the point of discrepancy to be investigated, (2) new information about actual performance and practice, and (3) a decision based on comparison of information and standard.

At each stage the program unit has four alternatives for further program activities: (1) go on to the next stage, (2) recycle the stage after there has been a change in the program's standards or operations, (3) recycle to the first stage, or (4) terminate the project. The evaluation staff helps identify gaps or problems, but the program staff must decide what to do about them.

SOURCE

Provus, Malcolm. *Discrepancy Evaluation*. Berkeley, California: McCutchan Publishing Co., 1971.

1.4 PRIORITY DECISIONS

ORIGIN	Cooperative Extension.
ESSENCE	This approach groups criteria for decision making about program priorities into six broad categories: <i>society-community, clientele, political, organizational, resources, and personal.</i>
USE	Although designed to be used in evaluating needs and alternative ideas for programs, it can also serve as a framework for examining the value and importance of completed programs.
STARTING POINT	Identify which of the categories are most important as criteria sources for the particular program.
DESCRIPTION	Program development requires decisions. These decisions are often difficult. Programmers usually are faced with too many problems to work on, too much content to teach, too many clientele groups to reach with the time and resources available. Therefore, priorities must be set. Criteria are needed for comparing and selecting from among program alternatives. The specific criteria to be used must evolve from the key influences affecting the particular program and programming situation. However, there are at least six general criteria areas that may need to be considered in establishing priorities. They include: Society-Community. The larger society, with its institutions, governmental units, pressure groups, trends, goals, and values, provides a broad framework in which programming must be carried out. The specific community also is considered with its economic, social, environmental, and cultural needs and desires; its formal and informal power and social structure; its linkages among various organizations; its potential for change based on past traditions and values. For example, the socioeconomic status of people needs to be considered--are there large numbers of low-income people? Also, is it politically feasible to shift resources to low-income programs? Does your personal value system say it's important to do so? Are other community resources being used for low-income programs? Clientele. There's a long tradition of basing programs on clientele needs. We all believe in and support this tradition. However, the question is: What clientele and what clientele needs? We need to look at clientele in terms

of subgroups within a community. Maybe we use Maslow's theory of needs as a framework for identifying priorities of various socioeconomic groups in the community. This could say to us that the focus for lower socioeconomic groups would be on physiological and social needs like nutrition or housing, while the higher socioeconomic groupings would focus on the upper part of the hierarchy--self-actualization.

Political. Programming decisions are made within political structures. We recognize the influence of government at the town, county, state, and federal levels in establishing program priorities. Governmental units should be used in our decision-making process, but not become such a constraining force that they overbalance all the other categories. There are times when program decisions aren't politically attractive, but we need to take risks and accept responsibility for influencing the political structure at various levels so more rational decisions can be made. However, the question: "Is it politically feasible?" is always relevant.

Organizational. Organizational statements of philosophy and mission provide insights to many questions about priorities for programs and clientele. For example: Is the organization supportive of innovations? What's the organization's position in programming with controversial issues like family planning? What kinds of programs does the organization reward? How much flexibility and personal input does the individual staff member have? What is the organization's position on programming with and for institutions and agencies that affect families versus programming with individuals or groups of individuals?

Resources. There are many questions about resources that need to be answered in setting priorities: Do we have the quality and quantity of resources necessary to effect change through a program? Are they the right kind of resources? Are we employing new personnel to coincide with changing program priorities?

Personal. We're interested in facilitating a self-perception among staff that they're educational leaders. If we support this concept, then provisions for use of personal values and perceptions must be allowed in the decision-making process.

The relationship among the categories must be considered. For example, you have to consider the political feasibility along with personal desires and needs of clientele in making a specific decision about programming with a disadvantaged group. The categories can be viewed as constraining and driving forces interacting against one another in a decision-making process. The third way of viewing the use of these

categories is as "screens" at which point a decision may be made to go or not to go with a specific program because of a judgment based on criteria within one of the categories.

SOURCE

Boyle, Patrick G. "Criteria for Program Priorities."
Paper presented at a conference of Extension Home
Economists Program Leaders, Washington, D.C., November,
1972.

1.5 DEVELOPMENTAL EVALUATION

ORIGIN	Development of curriculum packages and prototype programs.
ESSENCE	Both the original design and the design as implemented must be carefully evaluated with immediate feedback into the design.
USE	Particularly useful to those developing pilot programs or packaged curriculum. However, also important in evaluating the "first run" of a program that will be repeated many times.
STARTING POINT	Determine the purposes of evaluation and how evaluation fits into the total development process.
DESCRIPTION	The first example described below comes from elementary and secondary education (Individually Prescribed Instruction Program [IPI] of the University of Pittsburgh Learning Research and Development Center). The second comes from the Canadian NewStart adult education program. <i>IPI FORMATIVE EVALUATION</i> Formative evaluation is the continual evaluation of all elements of a developing educational program as an aid to the development process. It's concerned with this basic question: How can every element and operation in the program be examined so it contributes to its improvement? It has four sub-questions: 1. What <i>goals</i> should the program achieve? 2. What is the <i>plan</i> for achieving these goals? 3. Does the <i>operating program</i> represent a true implementation of the plan? 4. Does the <i>program</i>, when developed and put into operation, achieve the desired goals? The four questions identify sequential steps in planning and developing a program. They must not only be examined in the evaluation, but also in relation to one another. Questions like the following are examined within the four steps.

Goals

1. Are the statements actually goals?
2. Are the stated objectives the real goals of the program?
3. Are the goals worthwhile?
4. Are the goals attainable?

Plan of the Program

1. Does the plan give promise of contributing to the achievement of the goals?
2. Is the plan developed in sufficient detail?
3. Will the plans and procedures be easily understood by the people who are to put them into operation?
4. Is it probable that the plan can be carried out?

Program Operation

1. What are the specific points to observe in operation analysis? According to the plan, what should the students be doing? What should the teachers be doing? What other personnel are involved and what observations should be made of their activities?
2. Are the activities actually being carried out according to the plan?
3. How can the operation be made to correspond more closely to the plan?
4. Does a study of the actual operation suggest any modifications in the plan? The feasibility and effectiveness of a plan can only be determined after it has been studied in operation.

Assessment of Results

1. Do plans provide for assessing results?
2. Are the assessment procedures reliable?
3. Is the total evaluation process comprehensive enough to provide the needed total picture of the program results?

4. What are the implications of the results for modifying the program?

Formative evaluation ensures that each step in program development is done with care and assesses the quality at each step. Weaknesses discovered at any step have implications for modifying the proceeding level. Feedback is essential in formative evaluation.

IPI evaluation also includes:

Individual Pupil Monitoring. The regular and systematic evaluation of student achievement for adapting instruction to individual needs.

Summative Evaluation. The evaluation of the results produced by an educational program for purposes of making judgments concerning its value.

SOURCE: Lindvall, C. M., and Richard Cox. *The IPI Evaluation Program*. AERA Monograph Series on Curriculum Evaluation. Chicago: Rand McNally & Company, 1970.

**NEWSTART EVAL-
UATION SYSTEM**

The primary purposes for developing an evaluation system are: (1) to get information required so that course developers can make appropriate decisions to accept, modify, or reject developed components and concepts (formative evaluation), (2) to get information that's required for dissemination activities (summative evaluation), and (3) to get information that will help the users in the installation of developed and tested course packages.

The general scope of evaluation is indicated by: *topic areas where decisions are made, type of data base used in the decision, source of the data for the decision, and time of the decision.*

Areas Where Decisions Made. Decisions are made about policy, level of intervention, type of intervention, areas and agencies of intervention, dissemination practices, organizational structure, methods and systems of evaluation and measurement, theory and models, training methods, student progress, coach characteristics (ability, training, personality, style), course/curriculum content, and so on.

Data Base. The type of data base on which these decisions are based can be divided into data from the *implementers of interventions* including such things as objectives of agencies;

social pressures, personal values, and prejudices of policy makers; professional judgments and opinions; priorities and standards; etc.; and data from the *people receiving the intervention* including such things as attitude and personality measurement, statistical results (descriptive statistics, difference tests, measures of association, etc.), standardized tests, skill levels of students, student employment rates and incomes, demographic statistics on students and potential students, and so on.

Source of the Data. Is similar to the data base, but involves more detail about the people on which the data are gathered, for example, funders, support organizations, communities, user agencies, present and potential students or clients, outside professionals and consultants, outside evaluators, project personnel (management, development, training, research, and support staff), and sometimes even abstract logical and rational considerations.

Time of the Decision. This can range from very early policy decisions through the various stages of development of an intervention or program to decisions about where to implement programs. All of these aspects of the decision evaluation process must be considered in formulating a complete evaluation system.

Formative/summative evaluation integrates different methods into a total research study: direct observations create a global picture, surveys and standardized tests supply objective data to identify individual and group differences, case studies offer a connected sequence of events to help determine and explain individual and group changes, the experiment provides a controlled test of the effectiveness of specific variables, and the overall research evaluation design provides for an interweaving of the various methods.

A more specific modeling of the activity of evaluation includes:

1. *Determine decision areas*--both decisions that must be made and those that already have been made.
2. *Determine who the decision makers are.*
3. *Determine information requirements*--what information is required by when.
4. *Determine methodology*--source of data, instruments, how obtain, how process, how report.
5. *Prepare detailed procedures.*
6. *Get approval of component development.*

The model charts: (1) the kinds of evaluation information needed, (2) the stage where the information has primary or secondary use, and (3) the type of information--whether it's logical analysis, professional judgment, opinions, etc., or whether it's gained from empirical trials and measurement of effects on students.

The following questions illustrate the kinds of information needed:

1. How effective or efficient is the master or prime instructional strategy (individual prescribed instruction, group process)?
2. How closely does the developed course reflect the concept expressed in the outline design specification?
3. How closely does the developed course fit the parameters and bounds described in the outline design specification?
4. How valid and complete is the course content?
5. How closely do the instructional objectives meet the stated criteria?
6. How realistic are the stated student entry characteristics?
7. Lesson plans/curriculum guide and associated methods and materials:
 - a. How closely do they meet the stated criteria?
 - b. How closely do they result in student achievement of the instructional objectives?
 - c. How effective and efficient is the student-instructor organization (specialist, coaches, one coach per group)?
 - d. What additional or alternative methods or materials exist for achieving the instructional objectives?
 - e. How effective and efficient is the student-coach ratio?

8. What's the optimal sequence in each course and between associated courses?
9. What effect does group composition have on outcomes?
10. To what extent do the students achieve the instructional and course objectives?
11. To what extent do the students retain and apply the skills, knowledge, and attitudes learned?
12. How compatible are the instructional staff behaviors with the course objectives?
13. How appropriate are the standardized tests used for the student population?
14. How well do the administrative and support systems meet the needs of the potential users of the course package?

SOURCE: Lamrock, A. L., A. D. Smith, and P. W. Warren.
"Evaluation: Its Scope and Systems for Evaluation Development." Paper prepared for Meeting of Research Directors of NewStart Corporation, Ottawa, Canada, March 29-31, 1971.

1.6 MATERIALS EVALUATION

ORIGIN	Evaluation of new curriculum guides and audio-visual materials.
ESSENCE	Decisions among materials should be made using a firm set of criteria and an orderly process.
USE	Used whenever you're selecting a packaged curriculum or an aid to use in a program.
STARTING PROGRAM	Determine the qualities needed in the item and the range of available products that may meet the need.

DESCRIPTION

Although data gotten through questionnaires and interviews with teachers and students who are using the material is important input into the judgment, this approach depends primarily on professional judgment. Using a model of some type helps you identify the key aspects where comparisons must be made and avoid being swayed by highly motivating aspects of a particular program.

"Trade-offs" must be considered carefully. A program may be strong in one area and weak in another. How does this compare with a second program that has different strengths and weaknesses? Which will you trade for the other?

Here are two examples of approaches to materials evaluation. One deals with it as a process. The other outlines and weighs criteria to be applied.

TRADE-OFF AND COMPARATIVE COST APPROACH

One approach outlines the following activities:

1. *Describe* the product to be evaluated as completely as possible. Description includes such things as intended use, length, time involved in using, who the potential customers are.
2. *Evaluate the goals* of the product. Are the goals appropriate?
3. *Clarify the point of entry* of the evaluator. Which decisions are irreversible? Which can be influenced by the evaluator?
4. *Determine the kinds of trade-offs* involved. What other ways are there of attaining the same goals? What's given up by accepting this particular product?

5. *Compare costs* of the product with costs of other alternatives. The comparison includes costs involved in handling, storing, repairing, and using as well as original cost of purchase.
6. *Make an intrinsic (secondary) evaluation.* Such evaluation might include technical quality, content evaluation, use of the uniqueness of the medium, feasibility of use.
7. *Make an outcome (primary) evaluation.* Such evaluation might include learning rate and knowledge acquisition and retention.
8. *Form judgments and recommendations.* Recommendations based on the judgments may need to be specifically directed to potential purchasers, producers, and funders.
9. *Stipulate circumstances* that might modify the conclusions.
10. *Evaluate the evaluator.* What might his biases be? What was over-emphasized? What was left out or given inadequate attention during the evaluation?

SOURCE: Glass, Gene V. "Educational Product Evaluation: A Prototype Format Applied." *Educational Researcher*, I (January, 1972), 7-10.

WEIGHTED CRITERIA APPROACH This technique calculates the cost-effectiveness of alternative curriculum materials by a detailed breakdown and analysis of their components, quality, and cost. The model includes the following components and weights:

- I. *Coverage* (50).
 - A. National, state, and college standards (30).
 - B. Qualitative comparison with standard materials (20).
 1. Scope (4).
 2. Detail (4).
 3. Accuracy (4).
 4. Clarity (4).
 5. Logic of sequence (4).

II. *Appropriateness* (50).

- A. Utility (10)--type of job preparation.
- B. Difficulty (10)--for teacher, for student.
- C. Suitability of media to material (10).
- D. Suitability of media to student needs (10).
 - 1. Direction (5).
 - 2. Structure (5).

III. *Motivational Effectiveness* (50).

- A. Teacher (20).
 - 1. Stimulation (6).
 - 2. Perceived utility (8).
 - 3. Demand level (6).
- B. Student (30)
 - 1. Stimulation (4).
 - 2. Perceived utility (10).
 - 3. Cultural relevance (14).
 - 4. Demand level (2).

IV. *Cost* (50).

- A. Dollar cost (30).
- B. Time cost (20).
 - 1. Student (10).
 - 2. Teacher (10).

When two or more alternative materials are being explored, the weights would be used in computing scores for each of the alternatives. The score would range from 0 to 200. However, in addition to comparing the total score, sub-ratings should also be considered.

SOURCE: Crane, Peter, and Clark C. Abt. "A Model for Curriculum Evaluation." *Educational Technology*, IX (October, 1969), 17-25.

1.7 PARTICIPANT REACTION APPROACHES

ORIGIN	Adult education and extension.
ESSENCE	A variety of techniques including steering committees, analysis of attendance and re-enrollment records, and end-of-session reaction forms are used to get participant reactions to programs.
USE	The information is used to modify programs to make them more attractive and valuable to participants.
STARTING POINT	Identify what you want participants to react to.
DESCRIPTION	Almost every programmer has his own approach to examining participant reaction. Rather than presenting any one model, an overview of the approach will be given. There are two general approaches to getting reactions. In the first and most commonly used approach, the programmer specifies what he wants information about. In the second, he uses an open-ended approach and gets the participant to comment on what the participant sees as most important in terms of how he reacts to a program. When the programmer sets the parameters on what will be reacted to, he usually concentrates on one or more of the following: 1. <i>Content</i>: newness, clarity, relevance, and importance. 2. <i>Format</i>: timing, adequate involvement, interest holding, ease of learning, understanding what's expected, and suitability of techniques and approaches used. 3. <i>Facilities</i>: location of meeting, room arrangement (see, hear, etc.), and auxiliary resources (meals, etc.). 4. <i>Teacher's Performance</i>: style of teaching, organization, ability to relate well to participants, hindering mannerisms, ability to communicate, skill in involving others, etc. 5. <i>Perception of Value</i>: what was learned, what will be used, how important or useful, was it worth the time to attend, were objectives met, were expectations met, did it satisfy the need, and were underlying assumptions sound?

6. *Follow-Up and Suggestions for Future Programs:*
what more is needed on this subject, and what other programs are needed?
7. *Variation in Reactions According to Selected Characteristics:*
experience, position, age, education level, first-timers, etc.

When participants are left to react to programs without any structured guidance, they'll respond either in terms of what bothers them or in terms of what they think they should say. It's useful sometimes to ask how they describe a "good" program and then ask them to rate the program on those characteristics. Often the major points that concern them are: (1) the value, practicality, and use of what they get from the program; (2) how participation makes them feel (motivated to do something, happy and worthwhile, etc.); and (3) extent to which their interest is held. Details in terms of content, performance, etc., are important only to the extent that they affect how the adult enjoys participating and his feeling of getting something worthwhile from that participation.

Systematically exploring participant reactions is especially important when you:

1. Are doing something new--new clientele, new content area, or new teaching approach. Getting kinds of feedback that will help you determine how to adapt and adjust the new becomes important.
2. Know your clientele so well that you take them for granted.
3. Have an uneasy feeling that things aren't going as well as they should, but you can't really put your finger on why.
4. Are doing something that you feel is especially important and want to know how others react.
5. Are doing something that you expect will be particularly difficult, conflict producing, valuable to some but not to others, etc., and you need to see how the total group reacts.
6. Know there's a split in feeling, but don't know the numbers or relative strengths involved.
7. Will be using the same resource people again and need evidence in hand to discuss changes in their contributions with them.

Although we're usually curious about how people react to programs, just getting reactions is of little value unless you're willing to listen to what's being said and use participant reactions in making further decisions.

SOURCES

Steele, Sara M. "Exploring Participants Views of Programs." Material prepared for students in a course on Evaluation in Extension, University of Wisconsin-Madison, Spring, 1972.

Steele, Sara M. "Determining the Quality, Importance and Suitability of Programs." Madison, Wisconsin: University of Wisconsin, Department of Agricultural and Extension Education, 1970.

ADDITIONAL REFERENCES

Kropp, Russell P., and Coolie Verner. "An Attitude-Scale Technique for Evaluating Meetings." *Adult Education*, VII (Summer, 1957), 212-15.

Knowles, Malcolm S. *The Modern Practice of Adult Education*. New York: Association Press, 1970, pp. 231-33.

Axford, Roger N. *Adult Education: The Open Door*. Scranton, Pennsylvania: International Textbook Company, 1969, pp. 195-204.

GROUP 2

EVALUATION OF PROGRAM PARTS

OVERVIEW:

What specific parts of a program should be evaluated? How do parts relate to one another and form systems? How do systems affect programs? Another way of patterning models for evaluation is to examine the program system and identify its key parts. This group of approaches provides a wide range of ways in which the program system can be examined.

EXAMPLES INCLUDED:

- An *execution-impact approach* for social policy makers to use in examining a program as a means of carrying out social policy.
- A framework developed through analyzing adult education and selecting *actual components*.
- A group of *managerial systems*.
- A group of *socio-organizational* approaches that emphasize human elements.

GENERAL ADVANTAGES:

- Programs are made up of things, activities, and people. Evaluation frameworks that identify the key elements, parts of a program, and their relationships are important in addition to frameworks that deal with stages and processes.
- Although the models differ in terms of what key elements they're talking about, all illustrate procedures in using a systems approach to evaluation.

GENERAL CAUTIONS:

- It's easier to analyze a system into its component parts than it is to put it back together again and evaluate the effectiveness of the relationships among the parts.
- Success rests with identifying the most crucial from among the many diverse elements in a program.

2.1 EXECUTION-IMPACT APPROACH

ORIGIN	Social science research.
ESSENCE	The two major elements of this approach are <i>assessment of execution</i> (finding out if the plan was carried out as it was expected to be carried out) and <i>assessment of the impact</i> of the program (what results it did produce).
USE	Although directed to social policy makers, it may be particularly helpful in examining programs carried out by many individuals or units working autonomously at different program locations.
STARTING POINT	Review the program plan.
DESCRIPTION	This approach is included in a description of how social science research can help social policy makers. The evaluation framework is part of a total framework that also calls for contributions of social research to the planning of the program. The need for good planning and careful design as a requisite for expecting to find much impact is an important aspect of the total model. In evaluation, the policy maker seeks to determine whether; 1. The program was carried out according to prescriptions set forth in the planning and development stages. 2. It worked. 3. The expenditure of resources has been efficient compared with alternative means of achieving the same objective. The evaluation section of the framework is concerned with execution and impact. Assessment of the Execution. This includes: 1. Determining whether the program has been directed at the appropriate target population. Eligibility criteria should be clearly stated and the program kept from bowing to pressures and diverting from the target clientele. 2. Keeping records on who's being served. This record keeping must not interfere with program operation or be an undue burden on resources.

3. Examining the extent to which the needs of the population are met.
4. Determining how well its actual operation conforms to the plan for the program. This may involve examining the effect of various styles programmers use in carrying out the same program activities.
5. Examining programs through direct observation by means of field studies.

Assessment of the execution of a program is an administrative task. It ensures that the program that's being tested actually was carried out as expected. It provides data for better preparing others to carry out similar programs.

Assessment of the Impact. Examines efficacy and efficiency.

Efficacy. Determines whether the input changed the target group or environment in the direction expected, and whether such a change would occur again if the program were repeated.

Efficiency. Refers to the improvement in terms of some measure of unit cost. Time, reduced need, and other similar factors may be used in calculating efficiency as well as monetary units.

When two programs or approaches to programs are being compared and both look equally effective, efficiency is a crucial determining factor.

SOURCE

Freeman, Howard E., and Clarence C. Sherwood. *Social Research and Social Policy*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1970.

2.2 ACTUAL COMPONENT APPROACH

ORIGIN	Analysis of adult basic education (ABE) programs nationwide.
ESSENCE	Evaluation is focused on: <i>staffing, recruitment, collaboration, in-service education, goals, instruction, and the relationship among them.</i> Each is seen as having inputs, processes, and outcomes. Intended practice for each component is compared with current practice to identify gaps for program improvement.
USE	Particularly useful in identifying crucial programming elements and monitoring and strengthening their functioning.
STARTING POINT	List the components of a program that are seen as crucial to the functioning of the program. Set the criteria for selecting those that will be evaluated.

DESCRIPTION

Two criteria were used in selecting the six components. More components can be added if desired. These six were selected because: (1) they showed ways ABE program people interacted interdependently to make the program function and (2) program people were most concerned about them.

Detailed guides were developed for getting information on intent. Current practice is determined through questionnaires developed for administrators, teachers, students, and co-sponsors as well as classroom observation forms.

Analyzing incongruities between intent and reported practice, the evaluation approach systematically examines qualitative differences in program expectations of administrators and supervisors, teachers, students, and among these three groups. Analysis of these data pinpoint real and potential trouble spots. Intent is also compared with current practice for each component and discrepancies in expectations among individuals and groups are identified.

Each of the six components is examined in terms of inputs, process, and outcomes. Specific aspects of the six components that were identified for examination are listed here.

<u>Component for Examination</u>	<u>Inputs</u>	<u>Processes</u>	<u>Outcomes</u>
<i>Staffing</i>	Administrative time Budget allocations Selection criteria	Searching Selecting Placing	Characteristics Competence

<u>Component for Examination</u>	<u>Inputs</u>	<u>Processes</u>	<u>Outcomes</u>
<i>Recruitment</i>	Administrative time Budget allocations Target population	People involved Methods used	Numbers Characteristics Interest
<i>Collaboration</i>	Administrative time	Working relationships	Recruitment Allies Facilities
<i>In-service education</i>	Money Time Number of staff	Methods used	Competence
<i>Goal setting</i>	Legal mandates	Deciding on priorities	Objectives Criteria
<i>Instruction</i>	Staffing Recruitment Collaboration In-service education Goal setting	Teaching- learning transaction Individualiza- tion Materials used	Persistence Achievement Further education Occupational development

The components are closely interrelated. Relationship is most apparent when the outcome of one component is a major input into another component such as when the outcome of recruitment is the number and characteristics of students--an important input into instruction. The central component is instruction. The other five components all are inputs into instruction.

This evaluation approach:

1. Is designed to help people compare what they're doing with what they want to be doing as a basis for making changes, so the results will be more satisfactory.
2. Is based on a naturalistic systems model in which evaluation activities are part of the feedback that facilitates a more effective relationship between inputs, processes, and outcomes to encourage ongoing organizational development.

3. Is organized to: (a) focus on an initial set of program components of concern to practitioners, (b) allow for the selection and modification of data collection instruments, and (c) facilitate the use of those instruments and procedures as examples in preparing evaluation materials for additional components.
4. Combines internal evaluation for familiarity and commitment to use results, with external evaluation for impartiality and awareness of outside resources and standards.
5. Includes procedures to incorporate current evaluation activities and data--such as enrollment, achievement, and financial reports to the state--and attendance reports that are so often used as indicators of teacher effectiveness.
6. Contains explanations of procedures designed to encourage people to participate in the evaluation activities and to use the results for program improvement.
7. Contains items that might be adapted for an exclusively external evaluation, even though it wasn't designed for this purpose.

Knox, Alan, Jack Mezirow, and Gordon Darkenwald. *Evaluating Adult Basic Education: A Strategy and a Manual*. New York: Columbia University, Teachers College, Center for Adult Education, 1973 [pre-publication draft].

2.3 MANAGERIAL SYSTEMS

ORIGIN	Industry and the military.
ESSENCE	Evaluation models drawn from management theory and practice give greater emphasis to the role of management in successful programming. They deal with system and interaction in the system as they would in getting products produced in industry.
USE	Particularly useful in more efficiently managing programs.
STARTING POINT	Analyze the system. Determine its components and their relationships.
DESCRIPTION	Administrative decisions are made on the basis of the judgment of the administrator and/or his consultant. Previous experience and a logical mind play a key role in these decisions. They are often instant and static. Managerial decision making is also based on the judgment, experience, and logical mind of the decision maker, but is guided by data that are systematically collected and analyzed. Management is the process of planning, operation control, supervision, and production engineering in which the decisions are data based. Two of the broad purposes for evaluation interface with the management concept: 1. To provide quantitative information or systematically collected data for consideration in decisions leading to the better conduct of an operation or to increase functional utility of a product under development. 2. To provide information or systematically collected data for consideration in decisions dealing with the adoption or support of a product or activity. There are many relevant frameworks in managerial science and operations research. A few have been included as examples to indicate how systems are analyzed and how such analysis is used. PROGRAM EVALUATION AND REVIEW TECHNIQUE (PERT) This technique, often known by its key initials PERT, involves a process by which: 1. Projects are broken into specific task components. 2. Specific tasks are graphically presented so that sequence and interface are clearly identified.

3. Responsibility for completing each task is clearly established.
4. Time estimations and deadlines are established.
5. Key steps and pivotal activities are identified.

The tasks, relationships, people responsible, and deadlines are diagrammed in flow chart form. Diagrams may be simple straight lines if only sequential relationships exist. However, most involve branching to show the sub-activities and other related activities. A flow chart is developed by starting with the deadline for the completed task and then working backward until all of the tasks have been charted. Three items are usually indicated: optimistic (O) time of completion, most likely time (M), and pessimistic (P) time.

The following criteria help in deciding whether to use PERT:

1. Does a specified objective exist? Can the accomplishment of this objective be determined?
2. Must some schedule date or deadline be met?
3. What's the degree of project complexity? As the complexity of a project increases, the need for PERT grows.
4. Does a degree of uncertainty exist about the definition of some of the program elements? It may be particularly helpful in complex situations; it's less apt to be needed in routine and standardized situations.

SOURCES: Roman, Daniel D. "The PERT System: An Appraisal of Program Evaluation Review Technique." In *Program Evaluation in the Health Fields*, Herbert C. Schulberg, Alan Sheldon, and Frank Baker, eds. New York: Behavioral Publications, Inc., 1969, pp. 243-53.

Ryden, Einar R. "PERT? What Is It?" *Keeping Current*, Washington, D.C.: Federal Extension Service-U. S. Department of Agriculture, March, 1967.

Cook, Desmond L. *Program Evaluation and Review Technique: Application in Education*. Washington, D.C.: U.S. Government Printing Office, 1966.

*ORGANIZATION AS
A TOTAL SYSTEM*

This approach suggests that managerial technology can be best used when the organization has been designed as a total system. The whole needs to be analyzed into its component parts.

Organization Is a Process with Inputs and Outputs. In business, the input is resources and the output is economic welfare.

Management Represents the Control or Feedback Mechanism. Given an expected output, the actual output is measured against it and, if a difference exists, a problem is indicated.

The management or control unit is separated into a series of parts to provide adaptive capability. One type of adaptive capability is the ability to change inputs to maintain the expected output.

The adaptor unit within the control system has an input analyzer (for example, market analysis) that indicates change and an identifier that indicates the state of the organization or any of its subprocesses at a given time (extent to which resources are presently fully used).

Both the input analyzer and the identifier feed information to the decision maker. He applies a set of rules that enables management to make decisions about changes in input, output, and/or process to adapt to changing needs and conditions.

The organization is apt to have several major subprocesses. In industry, they include marketing, production, finance, and personnel. Each has its own adaptor. Each is sensitive to and reacts to both the other subprocesses and external stimuli.

Learning capability in the form of a designer is added to the adaptive system. The learning capability is the ability of system to redesign itself and learn from past mistakes to improve system performance.

The *design process* includes: (1) the originating problem, (2) a model, (3) specifications, (4) pilot simulation studies, and (5) field operations. The model includes the following elements: input, output, process, system logic, and information requirements.

Inputs are of three kinds: the input that enters the system and on which the process is to operate, environmental inputs that affect the process, and instrumental or control inputs that modify the operation of the process or the process itself.

Design criteria are rules used to evaluate the acceptability of designs. The most widely used are measurability, feasibility, optimality, reliability, and stability. Measurability is the system's ability to evaluate its performance. Feasibility indicates whether the model will operate as planned. All parts of the system must be integrated and mutually consistent for feasibility to occur. Optimality is concerned with the best choice. However, the optimal choice may not be feasible. Something that will work comes first and then optimality is dealt with.

SOURCE: Young, Stanley. "Organization As a Total System." In *Emerging Patterns of Administrative Accountability*, Lesley H. Browder, Jr., ed. Berkeley, California: McCutchan Publishing Corporation, 1971, pp. 311-30.

**MACRO SYSTEM
MODEL**

Evaluation is defined as the process of first identifying and then quantifying the relationships between student inputs and educational outputs, given a constant financial input and controlling for effects of external systems.

Student Inputs. Refer to the nature and characteristics of the students entering the program to be evaluated.

Educational Outputs. Include (1) changes in students as a result of the program and (2) impact of the program on home, community, other programs, etc.

Financial Inputs. Refer to the financial resources available to carry out the program.

Mediating Factors. Are the descriptive characteristics (personnel, organization, instructional design) of the way the financial inputs are used in the program in combination with the student inputs.

External Systems. Mean the social, political, legal, economic, and other systems outside of the school, formal and informal, that encompass the program, have impact on it, and are in turn modified by the outputs of the program.

The model is pictured with student and financial inputs entering the system, being exposed to the mediating factors and then exiting as student and other outputs. This occurs within the context of the effects of external systems.

Student outputs are seen as more than scores of students on academic achievement tests. They include changes in attitude and practice and the students ability to deal with real-life situations.

The changes that occur through education have social, political, and economic implications. The very nature of the external systems is altered by changes in student outputs.

SOURCE: Alkin, Marvin C. *Towards an Evaluation Model: A Systems Approach*. Working Paper No. 4. Los Angeles: University of California, Center for the Study of Evaluation of Instructional Programs, August, 1967.

*SYSTEM APPROACH
TO GOAL SETTING* System analysis is a process, a rational mode of approaching complex problems, a tool that helps in sound decision making. Developed to help administrators better relate the systems approach to program budgeting, the framework has the following main parts:

1. Systems definition, structure, and design.
2. System design process.
3. Defining the boundaries of the system and identifying how those within the system judge the performance of that system.
4. Need for goal integration and method to obtain it.
5. Goal-setting process.

Key ideas within the parts include:

Systems Definition, Structure, and Design. A system is defined as a set of elements whose relationships are designed to accomplish a desired goal. This particular framework deals with the school system. The elements of the school system are identified as: *philosophy, goals, objectives; inputs and/or resources; outputs; the environment; the programs; the agents, decision makers, and management; and measures of effectiveness.*

Programs are defined as the processes by which inputs are converted into outputs. In the school context, programs are the alternative types of curriculum that could be devised to achieve the objectives. Programs can also be broken down into activities, projects, components, and/or elements.

The element of *agents, decision makers, and management* stresses the role of people in management. A management system is an organization and as such involves agents and decision makers that make it work. It's important in system analysis to account for all the decision makers across organizational boundaries who contribute to the performance of a particular function.

System Design Process. System design is a problem-solving process starting with the needs of the client. Needs are translated into specific goals and objectives that are expressed in terms of measures of effectiveness. These measures serve as standards of achievement against which the results of alternative courses of action can be compared. The evaluation of outcomes will be made in the context of a value model that represents the relationship among the important variables of the system. Results will be compared to the original design and a new level set dependent on the degree of success or failure achieved.

Defining the Boundaries of the System. Education means different things to different people. Teachers and parents may not have the same perceptions. Actors within the system who hold expectation about the system include: teachers, non-teachers, parents, students, community, nation, and university and higher education. Decisions are affected by how administrators set boundaries for inclusion or exclusion of people with concerns about education. Conflicts often occur. For example, teachers may be concerned with "best" education, while taxpayers want average education commensurate with reasonable taxes.

Goal Integration. The administrator must get general agreement on the general philosophy and broad goals from all the groups in the system. The goals and objectives, and the ability to attain them, are influenced continuously by the constraints on the system which, in turn, is influenced by the cultural environment it operates in.

Goal-Setting Process. The goal-setting process starts with a statement of general philosophy or mission, then moves through general goals, subgoals, and objectives. To be useful an objective must be (1) quantifiable or observable, (2) definite, (3) measurable, (4) obtainable, and (5) in conformation with the organization's goals.

SOURCE: Van Gigch, John P., and Richard E. Hill. *Using Systems Analysis to Implement Cost Effectiveness and Program Budgeting in Education*. Englewood Cliffs, New Jersey: Educational Technology Publications, 1971.

**PROGRAM CONTACT
SYSTEM**

A system approach is a managerial effort taken to initiate an operational program that's designed specifically to generate a smooth, effective, and efficient flow of information from those providing it to those in need of it for decision-making purposes.

Within the Extension program planning format, the system approach includes a structure, an interacting complex of Extension and research people, facilities, and procedures designed to generate an orderly flow of pertinent information for dissemination and use as the basis for decision making in specified responsibility areas. This information is collected from both intra- and extra-program sources.

The increasing volume of information flowing from university research units is likely to continue. If Extension is to realize its mission of extending university knowledge off campus, programs must be planned and organized to accommodate and facilitate this increased flow. Information must first be evaluated to see that it's: (1) suitable to the audience's needs, (2) sufficient in quantity and quality to help in problem solving, and (3) provided in time for the recipient to use it.

The following programming system is proposed:

Phase I--Program Formulation. Includes: (1) information external to the specific program, (2) research needs according to established priorities and tempered by current research capabilities, and (3) stored information awaiting the activation of a new program. In the beginning stages, program formulation may only consist of assembling a series of ideas, concepts, and policies.

Phase II--Decision Phase. This phase is made up of four subphases: (1) program development, (2) program strategy, (3) program dissemination strategy, and (4) program promotion strategy. Program development involves explicit recognition of audience characteristics and needs and consideration of alternative program approaches that will increase effectiveness. Program strategies consider the socioeconomic and institutional environments and restraints.

Dissemination strategy necessitates the study of: (1) existing channels of information distribution, (2) types of informational outlets, (3) size and general characteristics of the setting clients live in, and (4) capabilities of the disseminators. Promotional strategy includes media efforts to precede, accompany, and follow the program. All of the strategies would be combined into one workable package.

Phase III--Execution Phase.

Phase IV--Evaluation of Program Results. This must be a joint effort between researchers (source of information that's disseminated) and extendors. Historically research personnel have had no way of determining: (1) if their efforts proved of practical importance, (2) if their knowledge was extended

in a timely and efficient manner, (3) if their "laboratory findings" were valid (or invalid) in general application, and (4) if a redirection of their efforts or of those of Extension personnel were warranted.

SOURCE: Duft, Ken D. "Systems Planning for Extension."
Journal of Cooperative Extension, VII (Fall, 1969), 168-78.

MANAGEMENT
INFORMATION
SYSTEMS (SEMIS)

Cooperative Extension adopted a nationwide management information system a few years ago. The system plays a dual role in the conduct of

Extension programs.

SEMIS (State Extension Management Information System) consists of planning, activity reporting, and some type of evaluation. One part of SEMIS in absence of the other two is incomplete. However, each part when analyzed may offer help in decision making and/or resource allocation. When these data are used together, they make a base for administrative and professional decisions.

SEMIS provides the professional a system for categorizing planned effort and for evaluating the effectiveness and efficiency of those efforts. The two roles aren't mutually exclusive, but complement each other.

Each decision is based on the best information available, whether it be hearsay, observations by untrained individuals, or factual data collected through a formal system. Obviously, the latter source of data is superior. Therefore, the management function is concerned with the total efforts planned or expended to accomplish a purpose and the benefits derived from the Extension effort.

The data provided through a formal system such as SEMIS are provided by each person who planned the program, carried out the tasks, and evaluated the results of the work. He knows better than anyone else what he did, when he did it, how he did it. He also has the competency to evaluate its effectiveness.

SEMIS isn't designed nor is it capable of serving as a substitute for professional judgment. It's a way of recording and transmitting these judgments. It provides a means for categorizing planning, activity reports, and evaluation to be compatible with all other staff member's. SEMIS provides two things: (1) a factual record of when, how, where, and with whom each activity was planned and conducted and (2) the contributions the specific professional and his activity made to the overall program.

SOURCE: *EMIS Data Definition Document, Fiscal Year 1974.*
 Washington, D.C.: Extension Service-USDA,
 Extension Management Information System.

**EVALUATING
 DECISION
 MAKING**

One approach to evaluating a decision requires the construction of a decision matrix. The matrix consists of: (1) a problem requiring a decision, (2) a list of alternative decisions, and (3) a generic list of criteria used to consider each alternative decision.

The generic list often contains items classified into various types. Three classifications are: (1) *technological*--is the solution workable with the available resources, (2) *sociopsychological*--is the solution likely to maintain morale, and (3) *economic*--does it fit within the economic environment and policies that affect that environment? The generic criteria must be developed into specific operational criteria expressed in such terms as resource load, results, and value of results.

The cells of the matrix are weighted as either positive, negative, or neutral. Actual decision making is observed and the number and types of cells actually used are noted. For example, how many alternatives are considered? How many and which criteria are used?

Adequate analysis takes into account both the final decision itself and the strategy used.

SOURCE: Hesseling, P. *Strategy of Evaluation Research.*
 The Netherlands: Royal Van Corcun Ltd.,
 pp. 217-34, as cited in *A Strategy for
 Evaluation Design*, Casper F. Paulson, Jr., ed.
 Oregon State System of Higher Education, 1970,
 section VII, pp. 22-24.

**ADDITIONAL
 REFERENCE**

Lange, Robert R. "Evaluation As Management Methodology."
 In *A Strategy for Evaluation Design*, Casper F. Paulson, Jr., ed.
 Oregon State System of Higher Education, 1970, section VII.

2.4 SOCIO-ORGANIZATIONAL SYSTEMS

ORIGIN	Social science theory.
ESSENCE	These models focus on the human rather than the non-human components of programming. They're concerned with facilitating changes that will improve programming.
USE	Programs aren't done by machines. Success lies with the people involved. Models that examine social and psychological dynamics give good information on how programs can be improved.
STARTING POINT	Identification of the human and social system elements that are most relevant to the success of programs.
DESCRIPTION	The quality of the program and the extent of its accomplishments depends on the quality, organization, and capability of the agency that offers it.

The purpose of evaluation isn't to prove but to improve and to facilitate activities necessary for improvement. Changes included in new programs and changes in old programs that are recommended by evaluation reports have to be made by people. Bringing about such changes needs as careful planning as does the actual programming or its evaluation. Change can be facilitated by evaluative processes that take adequate advantage of the system and factions involved.

There are several social and social-psychological models that might be adaptable to program evaluation. Attention is just beginning to turn to these approaches and few of them have been thoroughly adapted and tested. Three models developed for other purposes are included as examples of types of frameworks that have potential for evaluating and improving programs.

SOCIAL SYSTEMS MODELS

Although designed for analyzing a social system, social systems may have a good deal to contribute to evaluating how an agency programs successfully. One of the most widely known models is that proposed by Loomis. It has three key kinds of ingredients: (1) *elements*, (2) *processes*, and (3) *requisites for social actions*.

Elements. Include ends or objectives, norms or standards, facilities, power, rank, sanction, status roles, knowledge, and attitudes.

Processes. Include communications, socialization, institutionalization, boundary maintenance, systemic linkage, and social control.

Requisites for Action. Include time, territoriality, and space.

Although you may examine the same ingredients, the relationships that are brought into focus in program evaluation may be different. For example, prime attention would be given to the elements of standards and norms and ends and objectives as envisioned by a social system (programming agency, government bureau, or cliente group) with the relationship of power, rank, sanction, and status and role examined in relation to the other two elements. The effect of the social system processes both on the program being evaluated and on the evaluation need to be considered. For example, communication, boundary maintenance, systemic linkage, social control, and institutionalization may all affect both the program and its evaluation.

SOURCE: Loomis, C. P. *Social Systems: Essays on Their Persistence and Changes*. Princeton, New Jersey: Van Nostrand, 1960.

ORGANIZATIONAL MODELS

When applied in evaluation, the starting point isn't the mission and goal of the organization, but a working model of a social unit that's capable of achieving a goal. It's a model of a multifunctional unit and includes both the functions direct to the program and the maintenance and support functions essential to the system.

It deals with such questions as how close does the organizational allocations of resources approach an optimum distribution under these conditions and with this mission? Either a lack or an overabundance of resources in relation to given thrusts may be dysfunctional.

It also concerns itself with feedback mechanisms.

SOURCES: Etzioni, Amitai. "Two Approaches to Organizational Analysis: A Critique and a Suggestion." In *Program Evaluation in the Health Fields*, Herbert C. Schulberg, Alan Sheldon, and Frank Baker, eds. New York: Behavioral Publications, Inc., 1969, pp. 101-20.

Schulberg, Herbert C., and Frank Baker. "Program Evaluation Models and the Implementation of Research Findings." In *Program Evaluation in the Health Fields*, Herbert C. Schulberg, Alan Sheldon, and Frank Baker, eds. New York: Behavioral Publications, Inc., 1969, pp. 562-72.

*MOTIVATION
MODEL*

Motivation helps to explain human behavior. In education, concern is both with motivation within (professional) and the motivation without (the layman or client). In each instance there is *input, motivational processes, and output*.

Input. Input for both the client and the professional include: needs, goals, value systems, and abilities.

Output. For the client, it's participation and behavioral change. For the professional, it includes effectiveness, efficiency, integration, and morale.

Motivational Processes. These, which convert the input into output, have been grouped into three categories: *organization, group integration, and personal*. Organization includes philosophy, rewards, and training; group integration includes goal setting, leadership, leaders/group members relationships, and formal and informal groups. Within personal is: achievement, recognition, communication, responsibility, and participation.

The model identifies evaluation criteria to be used related to the various processes. The following are selected examples:

Organization: Philosophy. Does the organization recognize the existence of individual's subconscious motives that influence actions? Does the organization strive toward helping people develop adequate self-esteem by treating the person as an individual?

Integration: Goal Setting. Are the goals clear, precise, realistic, and internalized? Have individual goals been established and are they in line with the organization's goals of efficiency and effectiveness?

Integration: Formal and Informal Groups. Are the norms and purposes of the informal group appropriate for accomplishment of the organization's and individual goals?

Reward: Achievement. Is achievement and level of aspiration of the individual recognized as a vital aspect of motivation? Does the organization provide outlets for the individual's pride of accomplishment?

SOURCE: Lewis, Robert B. "Motivation Model for Extension." *Journal of Extension*, X (Winter, 1972), 23-24.

GROUP 3

EVALUATION--KINDS OF DATA; TYPES OF ACTIVITIES

OVERVIEW:

What kinds of data do you deal with in evaluation? What types of activities are involved? This group of approaches "model" the kinds of data used in evaluation or the kinds of evaluative activities that can be employed.

EXAMPLES INCLUDED:

- The *countenance of evaluation* identifies three major categories of data about programs and four types of data that can be used in forming conclusions.
- The *system role model* categorizes data in terms of program elements.
- The *means-ends hierarchy* shows how various kinds of common evaluative data on process and results relate to each other.
- The *when-to-do-it-yourself continuum* systematizes the kinds of data-gathering activities usually carried on by program units and suggests when the teacher-programmer may need to involve an expert.

GENERAL ADVANTAGES:

- Understanding kinds of data and data-gathering activities helps you better understand evaluation.
- Such understanding also helps you be more selective in investing resources in getting data and managing data activities more efficiently.

GENERAL CAUTIONS:

- Data always have to be related to the purpose of the evaluation. There are many kinds of data that could be collected, but not all of them are crucial to each evaluation.
- The average programmer can't keep up on all types of data-gathering techniques; therefore, knowing when to call for help and where to turn is important.

3.1 COUNTENANCE OF EVALUATION

ORIGIN	Public school education.
ESSENCE	Evaluation data are categorized as being either <i>descriptive--intents and observations--or judgmental</i> . Both descriptive and judgmental data are gathered about <i>antecedents, transactions, and outcomes</i> . Examining <i>contingencies</i> and <i>congruencies</i> among the sets of data is an important part of evaluation.
USE	Particularly helpful in understanding program dynamics and in selecting the most crucial data for a particular evaluative purpose.
STARTING POINT	Develop the rationale for evaluation. Why is it being done? Who will use it? For what purpose?
DESCRIPTION	Education is best improved through adequately dealing with its complexity and dynamic nature. The components of the program are parts of an interdependent system. An evaluation model should identify the components of evaluation. Although these components may in some cases be the same as program components, the evaluation model considers them only as they're a component of the evaluation. The evaluation should focus on the parts of the program that will produce the most program improvement for the evaluation dollar. Educational evaluation has both an informal and formal side. The informal includes casual observations, implicit goals, intuitive norms, and subjective judgments. The formal includes check lists, structured visitation by peers, controlled comparisons, and standardized testing. Neither gives full accord to the dynamics of education. Evaluation includes two acts--<i>description</i> and <i>judgment</i>--both are essential. It seeks generalizations about educational practices. Even though an outside evaluator may be unwilling to judge, he must include both judgmental and descriptive data in his work. The kinds of data to be included are viewed in a matrix (two-dimensional table) type of arrangement. Horizontal Side. The horizontal side of the matrix includes the three major categories of data to be examined: <i>antecedents, transactions, and outcomes</i>. These are generic labels that cover a variety of real-life items. <i>Antecedents.</i> Are conditions existing before the teaching that may affect the outcome. For example, the status of the

student at the start--his past experience, aptitude, interest, level of knowledge, and skill--is an important antecedent. The term may be equatable with input if input encompasses student and community inputs as well as the input of the educational system.

Transactions. Are encounters of the student with teachers, other students, authors of texts--the succession of engagements that make up the process of education.

Outcomes. Are evidence of the impact of instruction on students, teachers, and others; changes in abilities; achievements, attitudes, and aspirations of students; wear on the equipment; effect on the school; cost. Outcomes include both those evident and existing at the end of the session and applications, transfer, and relearning that occur later. Outcomes are the consequences of educating--immediate and long range, cognitive and connotative, personal and community-wide.

Vertical Side. This dimension of the matrix describes the two kinds of *descriptive data* and the two kinds of *judgmental data* that may be relevant to each of the horizontal categories.

Descriptive Data. Include:

Intents. Are descriptions of the planned-for environmental conditions, the planned-for teaching methods and content coverage, and intended student behavior. The collective descriptions should provide a priority listing of all the things that are intended to happen. Emphasis is given to what teaching as well as what learning is intended. Part of the information comes from the rationale for the program.

Observations. Are descriptions of what actually did occur related to the antecedents, transactions, and outcomes. This category of data can include direct observations. It also includes the realm of data about transactions and outcomes secured through the use of tests, check lists, and other instruments.

The two basic ways of using descriptive data are:
(1) looking for *contingencies* among antecedents, transactions, and outcomes and (2) looking for *congruency* between the intended and the observed.

Exploring the degree of congruence between intended and real (as measured and/or seen) and the contingencies between the antecedents, transactions, and outcomes creates a much better understanding of the effectiveness of plans and implementation.

Judgmental Data. Two approaches are: (1) *absolute standards* as reflected by personal judgments and (2) *relative standards* as reflected by characteristics of alternative programs.

Absolute Standards. Indicate levels set for antecedents, transactions, and outcomes. Data provide information on the judgment of the extent to which the standards are met. This evaluative activity involves identifying possible standards and weighing or assigning relative importance to each set of standards considering how much attention should be paid to each reference group (students, curriculum experts, etc.).

Relative Standards. Are similar to absolute standards, but use other programs as the base of identification. The base of comparison may be either the intended or the observational data. In the first case, you're using judgments about the program as it's planned. In the latter, you're using judgments about the program as it was actually carried on. Both types of standards may include several characteristics. As a result of judging the program antecedents, transactions, and/or outcomes, a composite rating of merit can be developed to use in making further decisions.

Use of descriptive data helps you understand programs, but unless some type of judgmental data is also used, conclusions about worth and value aren't present and the evaluation hasn't been completed.

Often only parts of the total countenance are used in a specific evaluation. The framework aids in deciding what categories and kind of information will be examined in the evaluation. For example:

1. Is it to be primarily descriptive, primarily judgmental, or both descriptive and judgmental?
2. Is it to emphasize the antecedent conditions, the transactions, or only the outcomes; or a combination of the three; or their functional relationships?
3. Is the evaluation to emphasize congruence between intended and observed?
4. Is it to focus on one program or to compare two or more programs?
5. Is it intended to help in deciding between programs or to help in developing or improving a program?

SOURCES

Stake, Robert E. "The Countenance of Educational Evaluation." *Teachers College Record*, LXVIII (1967), 523-40.

Stake, Robert E. "A Research Rationale for EPIE." *The EPIE Forum*, I (September, 1967).

Stake, Robert E., and Terry Denny. "Needed Concepts and Techniques for Utilizing More Fully the Potential of Evaluation." In *Educational Evaluation: New Roles and New Means*, Part II, 68th Yearbook of the National Society for the Study of Education, Ralph W. Tyler, ed. Chicago: The University of Chicago Press, 1969, chapter XVI, pp. 370-90.

Stake, Robert E. "Objectives, Priorities and Other Judgment Data." *Review of Educational Research*, XL (April, 1970), 181-212.

3.2 SYSTEM ROLE MODEL

ORIGIN	Adult education.
ESSENCE	Key components of the model are: <i>context, inputs, process, activity, outcome, judgment, and application of findings.</i>
USE	Particularly helpful in identifying the range of specific data that might be examined about a program so that the most relevant can be selected.
STARTING POINT	Describe the program as a system. Identify the roles played within it. Determine the relationship of evaluation to the system and the roles.
DESCRIPTION	The general purpose of evaluation is to improve the educational program by facilitating judgments about its effectiveness based on evidence. Programmers and administrators should consider the following points as they design evaluation systems: Evidence. Systematic and continuous program evaluation procedures should provide more adequate evidence and improve the soundness of judgment. Benefits. The extensiveness of the evaluation procedures depends on the importance of making sounder judgments. There should be a balance between the costs of evaluation and the benefits received. Frequency. The frequency with which evaluation data is collected depends on the aspect of the program that is being assessed and the anticipated use of the results. Feedback. Evaluation's major function is providing continuous internal feedback to enable adjustments in the ongoing program. Commitment. Those who are affected by the evaluation should participate in the evaluation process so that the likelihood of their using the results will be increased. Objectivity. The outside evaluation specialist can help to increase the objectivity and validity of evaluation procedures. Standards. Appropriate standards for comparison should be established. Both the relative performance of similar programs and absolute standards of excellence should be considered.

Relevance. Data that are highly relevant to the intents and objectives should be collected and analyzed.

Values. The process of making judgments should include, in addition to data, appropriate values, consensus, and continuing commitment.

The system role approach includes seven components that are identified from system and role theory and the analysis of past research and evaluation. Each component has a specific set of evaluative activities.

Context. An understanding of the setting in which the program takes place and the influence of that setting on the program. Examining the program rationale is an important part of the setting. The rationale within which a continuing education program is developed and evaluated has three parts: history, current demands and constraints, and expectations about future development.

Five related activities include examining the setting, deciding on the most important evaluation emphasis, selecting evaluation models and procedures, and identifying the context in which the evaluation will occur.

Inputs. Include both what was intended and what actually occurred. Participants, materials, teachers, and administrators are all inputs into the program. Input characteristics of participants include:

1. Biographical--age, sex, travel experience, work history.
2. Status--prestige, level of work, income level, level of living.
3. Ability--performance on intelligence and aptitude tests.
4. Personality--attitudes, self-concept.
5. Roles--present family situation, type of employment.
6. Needs--gaps between present and desired circumstances.

Categories of information suggested for materials evaluation include:

1. Content goals--what knowledge should be gained?
2. Performance goals--what should the learner be able to do with what he has learned?
3. Teacher requirements--what should the teacher know to use materials?

4. Participant expectations--what should learners want to accomplish?
5. Community factors--characteristics of the community that influence effectiveness of material.
6. Activity--description of the intended and the achieved inputs into the programs.

Staff inputs should be considered in terms of roles--instructional, administrative, and support. Among characteristics to be examined are subject-matter competence and ability to relate to participants. Inputs that go beyond characteristics are support inputs like amount of encouragement and in-service education. Basic activities include determining the nature of important inputs and then describing the quantity of these inputs.

Examining inputs:

1. Makes it easier to plan to acquire the inputs.
2. Helps with the planning of intended learning activities and outcomes that can be achieved, based on the intended inputs.
3. Provides the basis for interpreting the learning activities and outcomes that actually occur.

Process. Description of the intended and the achieved transactions that bring together inputs to produce outcomes. The heart of the program is the teaching-learning process. Interactions in programming take place in four settings--individual, temporary group (classes), organizational (in-service), and community.

Intended processes are descriptions of the ways inputs should interact to produce outcomes. The achieved processes are those that actually occur. Plans are a major source of the intended processes. Information on what actually occurs can be gotten from analyzing interaction, using trained observers, rating scales completed by students, etc.

Activity. One way of analyzing process is how well it transforms resources to maintain stability and stimulate change. Examples of some of the things to be examined include: goals and policy, program development, teacher selection and supervision, learner selection, teaching-learning transaction, learner support and advisement, support staff selection, administrator selection, maintenance of personnel, adaptation and change, facilities and equipment, materials, coordination and communication, financing.

Outcomes. Vary from immediate stages that describe direct changes in learner's knowledge, skills, or attitudes to more remote stages where outcomes are stated in terms of benefits to the community. The key question is whether it's reasonable to expect these inputs and transactions to produce the intended outcomes.

Major categories of outcomes include:

1. General impact--progress participants make, proportion who complete, adequacy of progress.
2. Participant satisfaction--do students feel they have gotten enough?
3. Content mastery--test scores such as those on standard equivalency exams.
4. Personality--improved self-image and greater social awareness.
5. Work related--increased income, higher employment rate, reduction of welfare rolls.

Evaluative activities include describing the intended and the achieved outcomes of the educational program. Major sources of intended outcomes include expectations of participants, teachers, administrators, and policy makers associated with the program.

Comparison of intended outcomes and achieved outcomes (descriptions of both) may indicate that adjustments should be made to raise achievement, to adjust expectations, or both.

Judgments. Judgments about the extent the intended was achieved in terms of inputs, process, and outcomes have been made earlier as a part of analyzing the logical soundness of the program plan. Conclusion about needed changes in activities or expectations or both should have been formed.

The next stage in making judgments focuses on results. It involves internal and external comparison. Internal judgments involve the extent to which the actual inputs and processes contributed to the achievement of the outcomes and whether the outcomes (benefits) compare well with the inputs (costs). External comparisons analyze the results of the specific program against external standards.

Application of findings. Conclusions should be used to improve the program. Reports should focus on the questions and concerns that staff and others have, and be designed to result in application.

Guidelines for handling results include:

Validity. Evaluation procedures that are well planned and implemented to produce valid results increase the potential use of the findings.

Communications. The results of evaluation efforts should be communicated to those who can use them--learner, teacher, administrator, and policy maker--in a form they can readily understand.

Commitment. Involve those who can use the results in such a way that results are valid and there's commitment to use them.

Timing. Results should reach those who can use them during a time period in which their use is feasible.

Implications. The major implications of the findings should be included in both the evaluation report and the discussion of it.

Time. Ways should be found to allocate time for the study and use of evaluation findings.

Assistance. Technical help should be available both for additional analysis and interpretation of findings.

SOURCES

Knox, Alan B. *Program Evaluation in Adult Basic Education*. Tallahassee, Florida: Florida State University, Adult Education Research Information, Processing Center, 1971.

Knox, Alan B. "Continuous Program Evaluation." In *Administration of Continuing Education*, Nathan C. Shaw, ed. Washington, D.C.: National Association for Public School Education, 1969.

Knox, Alan B. "Improving Urban School Adult Education: An Analysis of the Adult and Community Education Programs of the Washington, D.C. Schools." New York: Columbia University, Teachers College, 1967.

3.3 MEANS-ENDS HIERARCHY

ORIGIN	Cooperative Extension.
ESSENCE	The model shows the relationship of seven levels of evaluation data: <i>inputs, activities, people involved, reactions, "KAS" changes, change in practices, and ultimate results.</i>
USE	Particularly helpful in situations where several kinds of input and result data can be generated and where the evaluator has to decide which to examine, or, if examining several, needs a rational way of relating them to each other.
STARTING POINT	Start with a decision about what information needs to be collected about the program, process, and program results. The model then helps you understand how this information relates to other information.
DESCRIPTION	The means-ends hierarchy deals with process and results. It's aimed at improving program development. It can be applied both in constructing agency-wide strategies and in designing specific evaluation projects. It can be applied through evaluative research, evaluative studies, and the "criteria" approach to evaluation. The three used in tandem is suggested as one systematic strategy for agency-wide evaluation. The model includes seven levels of evaluation: inputs, activities, people involved, reactions, "KAS" changes (knowledge, attitude, and skill), application of learning or "practice change," and results achieved, that is, degree of problem solution. Level I--Inputs. Description of the amount of effort, e.g., time and money spent. Simplest and easiest data to get. They're needed for cost-benefit analysis at a higher level. Level II--Activities. Descriptions of the kinds and frequencies of educational activities performed in the project. Level III--People Involved. Descriptions of numbers of participants reached and nature of balance in terms of social characteristics. Level IV--Reactions. Descriptions of what people say about the activities--whether they see they're valuable, their like or dislike of the quality and quantity of activities.

Level V--"KAS" Changes. Data on (1) whether changes have occurred in specific knowledge, attitudes, and skills; (2) the direction of such change; and (3) the distance or extent of change.

Level VI--Practice Change. Explores whether people have used their knowledge, skills, and attitudes and whether desirable patterns of behavior have resulted.

Level VII--Results Achieved. This category examines what consequences the changes in behavior have brought to people and communities. There's an underlying assumption that each of the lower levels do in fact have a bearing on bringing about the final results. The data in each of the seven levels can be related to each other. For example, inputs can be compared with "KAS" changes or with practice change or with results achieved. Or, recipient and reaction data can be compared with input or with results achieved.

Other propositions are:

1. Any evaluative question within the hierarchy is related to, and has implications for, all the rest of the questions.
2. Evaluation at the lowest levels of the hierarchy is rather meaningless for impact evaluation unless we either know or make assumptions about the answers to "higher" evaluative questions.
3. If a higher-level evaluative question is answered positively through research, the achievement of lower-level objectives can be interpreted as progress.
4. Evaluative studies at the higher levels should be stressed. These evaluation studies can help provide a basis for interpretation of less costly, lower types of evaluation.

SOURCE

Bennett, Claude F. *"Theory and Strategy for Evaluation of Knowledge Dissemination."* Washington, D.C.: Extension Service-USDA, Program and Staff Development, 1973.

3.4 WHEN-TO-DO-IT-YOURSELF CONTINUUM

ORIGIN	Cooperative Extension.
ESSENCE	Data-gathering activities are grouped into five main categories: <i>habitual but unorganized, simple guides, reporting by students or teacher, post- or pre- and post-testing, and experimental design using control or comparative group.</i> Although program personnel are responsible for carrying out most of the activities, help should be gotten from specialists for more complicated activities.
USE	Particularly useful in emphasizing the shared roles between program personnel and evaluation or studies specialists.
STARTING POINT	Identify the type of data activities needed in the evaluation.
DESCRIPTION	The major purpose of evaluation in education is to determine the effects of teaching under given conditions on the knowledge, attitudes, and behaviors of those being taught to provide a basis for improving, justifying, or discontinuing the teaching activity. The major focus of evaluation is to try to determine what kind of individual emerges from the learning experience to which he has been exposed. The continuum of types of data-gathering activities is as follows: <u>Habitual But Unorganized</u> 1. Habitual but unorganized introspections of the teacher relative to the teaching situation. 2. Panel or staff group who jointly and informally review a teaching situation without a list of questions or items to guide the review. <u>Simple Guides</u> 3. Teacher's rating by means of a list of questions or items of his teaching relative to what he did and effects on those being taught. 4. Observations of a teaching situation by an individual or by those taught, guided by a list of questions or items.

5. Observations of a teaching situation by a panel guided by a list of questions or items.

Reporting by Students or Teacher

6. A system used by the teacher for reporting the learning of individual students.
7. Analysis of an educational experience by students using a well-designed post-questionnaire.

Post- or Pre- and Post-testing

8. A post-questionnaire or test that determines the knowledge, attitudes, and skills of students resulting from one teaching situation.
9. A questionnaire or test that determines knowledge, attitudes, and skills of students resulting from one teaching situation.
10. A questionnaire or test before and after more than one teaching situation (often several as in a program extended over time) that determines the knowledge, attitudes, and skills of students resulting from the exposure; desirable to be accompanied by an input study.

Experimental Design Using Control or Comparative Group

11. A questionnaire or test before and after an educational experience (either one or more exposure) measures knowledge, attitudes, and skills; a similar pre- and post-testing of a matched control group; a study of the educational input.
12. Experimental study using a control or comparative group in which the educational experience (either one or more exposure) including methods is planned for studying what happens in students; before and after questionnaires or tests are used; a study of the educational input.

The framework suggests that the program personnel can handle Activities 1 and 2 by themselves. They may want some help from studies specialists for 3-5 depending on the complexity of the task and the programmers background. Type 6 activities are developed jointly by program personnel and studies specialist, conducted by program personnel. Activities 7-9 are developed jointly, but conducted by either program personnel or studies specialist. Activities 10-12 are planned and conducted by studies specialist with assistance from program personnel.

SOURCE

Alexander, Frank D. "A Critique of Evaluation." *Journal of Cooperative Extension*, III (Winter, 1965), 205-12.

ADDITIONAL
REFERENCES

Steele, Sara M. *Programmer's Responsibility in Examining Program Effectiveness*. Evaluation Series. Madison, Wisconsin: University of Wisconsin-Extension, Division of Program and Staff Development, October, 1972. [Identifies program personnel responsibility regardless of whether help is gotten from evaluation or research specialists.]

Kirkpatrick, Donald L. "Evaluation of Training." In *Training and Development Handbook*, Robert L. Craig and Lester R. Bittel, eds. New York: McGraw-Hill Book Company, 1967, pp. 87-112.

GROUP 4

EVALUATION PROCESSES

OVERVIEW:

How do you evaluate? What are the processes involved? Another way of "modeling" evaluation is in terms of processes. This group of approaches focuses on the processes of evaluation. The first three deal with basic processes used regardless of what's being evaluated. The other three present specific evaluation processes.

EXAMPLES INCLUDED:

- The *appraisal model* emphasizes the role of professional judgment.
- The *data management* framework provides a guide for dealing with quantifiable data regardless of what's being examined.
- The *natural process approach* shows how appraisal and quantitative approaches can be combined in aiding decision making.
- *Monitoring evaluation* shows how you can use evaluation during programming to keep the program on track.
- *Improvement evaluation* uses appraisal to stimulate program improvement.
- *Transactional evaluation* emphasizes the human relationships involved in facilitating change.

GENERAL ADVANTAGES:

- Although specific procedures and activities vary, evaluation relies on certain constant basic processes. Good understanding of these processes can speed evaluation.
- Processes that involve program personnel in a positive way are particularly important in getting evaluation used.

GENERAL CAUTIONS:

- Processes should be kept as simple and unobtrusive as possible.
- Simplicity should be viewed from the participants' eyes rather than from those of the evaluator.

4.1 APPRAISAL MODEL

ORIGIN	Accreditation bodies and educational agencies responsible for examining the programs of local units.
ESSENCE	This approach emphasizes <i>professional judgment</i> . Judgments are made by an expert, team of experts, or team of program peers--from within or without the system--who examine the program and form conclusions and recommendations. <i>Criteria</i> guides are used in examining the program.
USE	Particularly useful when immediate feedback and interpretation are needed and where interaction with program personnel facilitates the use of evaluation.
STARTING POINT	Identify the purpose of the appraisal and the purpose of the program that's being appraised.
DESCRIPTION	This process is used in accreditation, in reviews by state departments of public instruction, and in the comprehensive reviews of Cooperative Extension. <i>Appraisal is an act of judgment in which the judging implies both a criterion--a standard of some sort--and a pertinent description of what's being done.</i> The criterion and the observation must deal with the same thing. Appraisal involves the following <i>activities</i>: 1. Specifying the purpose of the appraisal. 2. Determining who will serve as the appraisers. 3. Establishing the purpose of the program. 4. Selecting or developing a set of criteria. 5. Identifying the aspect or aspects that will be evaluated. 6. Recognizing and understanding the implications of the assumptions that are being made when program aspects and criteria are chosen. 7. Amplifying the criteria so that they become a detailed statement of the kinds of observations that need to be made. 8. Developing a plan of action for making the observations (getting data a variety of ways). 9. Developing, modifying, selecting, and using techniques of observation.

10. Determining number of observations and procedures for increasing validity, reliability, and objectivity.
11. Recording, interpreting, and summarizing the observations.
12. Establishing bases to which the observations can be related (that is, selecting norms and standards as a base for interpretation).
13. Making the conclusions of the appraisal known.

The quality of the appraisal rests on the expertise of the appraisers and on the quality of the criteria used. *Criteria* constitute definitions of what is seen as "good." Criteria define what characterizes a good program. They express values or interpretations of what's good. Because values differ, there may be several criteria alternatives. What may be seen as good at one time and place may not be valued in the same way in another.

Criteria must be consistent with the educational philosophy that's being followed. They postulate values that should be achieved. They must be so formulated to permit making pertinent descriptions. They must be useable. This often means a greater precision in definition than is first thought.

Success of appraisal rests partly on answering the following criteria questions: What criteria? How defined? How applied? How limited? This model culminates in judgments and conclusions of people instead of, or in addition to, systematically produced data.

It's also important to understand the *assumptions* involved in the evaluation. There are usually at least two kinds:

1. Assumption of the purpose of education.
2. Assumption of correlation or relation.

The other key element is *observation*. Observations include both describing and interpreting what's observed.

The aspects of a program that are subject to appraisal fall into four categories:

1. Plans and purposes--organizational, administrative, and instructional objectives and plans.
2. Resources--material resources (such things as furnishings, and AV equipment) and personnel.
3. Processes--administrative, supervisory, and instructional.

4. Effects--results in terms of students, staff, and community.

Although appraisal is often thought of as primarily perception and interaction in on-site visits, it can also include review and interpretation of data about a program. The appraisal method is sometimes criticized for its lack of use of objective data. However, this may be more a fault of how it is operationalized rather than with appraisal itself. Data can be built into appraisal systems.

SOURCE

Harris, Chester W. "The Appraisal of a School--Problem for Study." *Journal of Educational Research*, XLI (November, 1947), 172-82.

ADDITIONAL REFERENCES

Herzog, Elizabeth. *Some Guidelines for Evaluative Research*. Washington, D.C.: U.S. Department of Health, Education, and Welfare, Social Security Administration, 1959. [Particularly good section on criteria.]

Checklists for Public School Adult Education Programs. Circular No. 356. Washington, D.C.: U.S. Office of Education, Federal Security Agency, n.d.

Honnold, Robert E. et al. "Comprehensive Reviews." Materials prepared by a Program and Staff Development Task Force, Extension Service-USDA, 1972.

4.2 DATA MANAGEMENT

ORIGIN	Educational research.
ESSENCE	The processes of getting, processing, and reporting the data used in evaluation have been presented in many ways. The framework accompanying the program decision model is particularly comprehensive in that it both outlines the crucial activities and categorizes some of the specific tasks where assumptions are often made that must be tested for validity in the evaluation setting.
USE	The focus on the validity of the assumptions made in the selection of sampling, treatment, measurement, and analysis procedures is particularly helpful.
STARTING POINT	Start with a clear analysis of the data needed for the evaluation.
DESCRIPTION	The major elements involved in <i>data management</i> include: (1) <i>delineating the information system</i>, (2) <i>obtaining the information</i>, (3) <i>providing the information</i>, and (4) <i>evaluating the evaluation</i>. Delineating the Information System. Includes defining the system being evaluated, specifying decisions inherent in the program, developing a statement of evaluation policies, and clearly outlining the assumptions upon which the study is based. Obtaining the Information. This second step includes collecting, organizing, and analyzing data. Three steps in collecting evaluation data are: (1) identifying sources of information, (2) developing the proper instrumentation, and (3) describing the conditions under which the data will be collected. The organization stage of obtaining the information consists of deciding on the unit of organization, storage and retrieval requirements, and quality control procedures followed after data collection. When the data have been collected and organized, the evaluation proceeds with the actual analysis. The first step is identifying the unit of analysis, which is determined by the level of the decision involved, the information required in making that decision, and the sampling unit employed. The basic tools of analysis include prediction, comparison, description, and explanation--the assumptions

of the approach employed determine the analysis, statistics, and reporting form used.

Providing the Information. Timing, frequency, and methods of these reports are determined by the audience for whom the evaluation is being prepared and the decisions within the program structure they relate to.

The first step in developing the report defines the audience it's intended for--decision makers, decision influencers, etc. This is relatively straightforward. But, the second task--depicting the reporting levels--is more difficult. Reporting levels can be dichotomized in terms of micro versus macro levels to define the level of detail of the information reported to the individuals within the programs and agencies. Consequently, evaluators must determine the audience for their reports, and accordingly produce them at the appropriate levels of generality or specificity.

A third aspect at this stage involves establishing the reporting setting, developing the content, and choosing the proper media to communicate the findings.

The information must be disseminated to the proper individuals. This involves, first of all, establishing channels for transmitting the reports and identifying the hierarchy to be followed in presenting the results. Reports should go to the administrator and the staff, before making the information available to funding agencies or the general public. A second consideration in disseminating reports is developing a procedure for publication. Abstracts may be provided to the appropriate news media, or where the information is unique, copyrighted to ensure protection over the use of the findings.

Six criteria should be kept in mind if evaluative data are to be informative to the receiver:

Relevance. The data are useless if they don't relate to the purpose they were collected for. The criterion of relevance asks whether the purposes actually have been achieved.

Importance. Data should be culled so that the most important information is emphasized. The person organizing the information must use what the user considers significant as well as his own view of what's important.

Scope. Data can be true, but not the whole truth. Scope ensures that weeding for perceived relevance and importance doesn't distort the data.

Credibility. Users must trust the source of the data. Credibility rests partly in the relationship between the data handler and the user of the data.

Timeliness. The information must reach the user when he needs it and/or is in a position to deal with it. Its value decreases if it's available too late or if it appears too soon.

Pervasiveness. Do all of the people who need to know about the evaluation get the information?

Evaluating the Evaluation. The final element in this model is applying evaluative tools to the evaluation study itself. Since the audience will be weighing the quality of his study, the prudent evaluator carefully examines his work against the criteria on which it will be judged.

Assumptions. The exploration of assumptions underlying various techniques and testing whether those assumptions are appropriate in the evaluation activity is an important element in designing evaluation. Four basic types of *assumptions* are of concern to evaluators: sampling, treatment, measurement, and analysis.

Sampling Assumptions. Include: (1) parameters of the population, (2) social structural characteristics of the population, (3) parameters to be tested, (4) objectives of the analysis, (5) representatives of the required sample, (6) precision which the sample affords because of the selection method employed, and (7) assumptions based on budgetary and resource limitations and precision relating to efficiency of sample selection methods for reducing uncertainty.

Treatment Assumptions. Include statements about internal and external validity. Internal validity assumptions relate to: (1) history of events between the pre- and post-test that may affect performance, (2) maturation of the subjects, (3) learning effects from the first test producing an impact on the results of the second, (4) instrumentation in which changes in calibration of measuring instrument or changes in the observers may produce changes in the obtained measurements, (5) effects of statistical regression, (6) biases from the selection of respondents for comparison groups, (7) experimental mortality, and (8) selection-maturation interaction in which designs are mistaken for the effects of the experimental variable.

External validity assumptions include: (1) congruence between population characteristics and those of the sample, (2) ecological validity, including both an explicit description of the independent variables assuming interference with subjects receiving two or more treatments, (3) Hawthorne effect, (4) novelty effects, (5) effect of the experimenter, (6) pre-test sensitization, (7) post-test sensitization, (8) interaction of history and treatment, (9) measurement of dependent variable required for generalization, and (10) interaction of the time of the measurement and treatment effects.

To establish limits of reliability in the treatment phase, the evaluator also should include assumptions about the additive effect of treatments, differential effect or treatment constancy, identical treatment of subjects, identical treatment of groups, random assignment of treatments, multiple treatments being avoided, and minimization of the interactive effects of the treatment with control variables.

Measurement Assumptions. Are another integral part of this stage of an evaluation design. They should deal with such basic questions as: (1) correspondence between sets of axioms about objects and numbers, (2) representativeness of the sample, (3) scale's compatibility with the phenomena being measured and analysis being employed, (4) validity of the instrument, (5) whether the scale of measurement is conceptually or operationally defined, (6) reliability of the instrument, and (7) objectivity of the instrument or technique. This information helps the evaluator control sources of error and gives the clientele of the evaluation a clearer picture of the limitations of the study.

Analytic Assumptions. State the intentions of the evaluation and outline the use the descriptive, comparative, explanatory, or predictive information will be put to. Analytic assumptions are concerned with the correspondence of statistical techniques with the properties of the data to be analyzed. Assumptions should be specified since statistical procedures are designed for or derived from them.

On the basis of assumptions about sampling, treatment, measurement, and analysis, the evaluator can develop a general picture of the kind of design or model that will organize and communicate these data most effectively. More specifically, these may include mathematical models, work breakdowns, PERT organizations, or an application of the critical path network technique.

SOURCE

Phi Delta Kappa National Study Committee on Evaluation.
Educational Evaluation and Decision Making. Itasca,
Illinois: F. E. Peacock Publishers, Inc., 1971,
chapter 6 and pp. 28-30.

ADDITIONAL
REFERENCES

- Caro, Francis G., ed. *Readings in Evaluation Research*.
New York: Russell Sage Foundation, 1971. [In particular,
see chapter 1 for an overview of the results of recent
experiences.]
- Corey, Stephen M. *Action Research To Improve School
Practices*. New York: Columbia University, Teachers
College, 1953. [Emphasizes action person doing research
in addition to the professional researcher.]
- Suchman, Edward A. *Evaluative Research*. New York: Russell
Sage Foundation, 1971.
- Weiss, Carol H. *Evaluation Research*. Englewood Cliffs,
New Jersey: Prentice-Hall, Inc., 1972.
- Freeman, Howard E., and Clarence C. Sherwood. *Social
Research and Social Policy*. Englewood Cliffs, New
Jersey: Prentice-Hall, Inc., 1970.
- American Institutes for Research. *Evaluative Research: Strategies
and Methods*. Pittsburgh, Pennsylvania: American Institutes
for Research, 1970.
- Gottman, John M., and Robert E. Clasen. *Evaluation in
Education*. Itasca, Illinois: F. E. Peacock Publishers,
Inc., 1972.
- Furst, Edward J. *Constructing Evaluation Instruments*.
New York: David McKay Co., Inc., 1964.
- Kaplan, Abraham. *The Conduct of Inquiry*. San Francisco:
Chandler Publishing Company, 1964.
[In particular, chapter on values.]

4.3 NATURAL PROCESS APPROACH

ORIGIN	Extension education.
ESSENCE	This attempt to provide a generic model of program evaluation that can be adjusted to evaluative needs emphasizes three major elements of evaluation: <i>criteria</i> , <i>evidence</i> , and <i>judgment</i> . It emphasizes improving the kind of evaluation that program personnel do naturally rather than replacing such activities with elaborate formalized procedures.
USE	Useful in improving ongoing everyday evaluation and in helping program personnel understand evaluation.
STARTING POINT	Start with how the individual usually evaluates. See how the accuracy and efficiency of the natural process can be improved.
DESCRIPTION	Evaluation assesses the value of programs and programming activities. There are three major elements in evaluation: <i>criteria</i>, <i>evidence</i>, and <i>judgment</i>. Evaluation takes place when judgments are made. Judgments are improved when they're made by comparing evidence about the aspect to be judged with criteria for what should be. Criteria. A criterion is a measure against which something can be judged. It may be a rule, a standard, a norm, an object or a condition, or behavior that's considered to be "good," "ideal," or of "high merit." It's a description or image of what a valuable (suitable, high quality, effective, important, and/or efficient) program is like. Criteria form the basis for interpreting information and developing judgments. They identify the values held by those involved. Sets of criteria exist for judging quality of meetings, workshops, and group interaction. There are fewer criteria statements that deal with overall value of an endeavor. Such statements are essential to setting priorities and assigning overall value to competing program requests. Criteria identify the aspects on which alternatives will be compared. Clearly understanding and trying to meet a reliable set of criteria for programming can do a great deal to improve programming efforts regardless of whether those criteria are used in any formal type of evaluation. Criteria tell us what we should be doing. Evidence. Evidence is an indication or an outward sign. In evaluation evidence is: acts, words, numbers, observations, or things that provide a sign or indication.

It provides proof of the extent to which the quality we're examining is present. It's the accumulated pattern of information that provides a picture adequate for judging the extent to which criteria have been met. Evidence is information accepted as descriptive of what's directly germane to the judgments involved. Use of research-type techniques may improve the quantity and quality of evidence. Regardless of the extensiveness of the data procedure, the evaluator must be able to assess the effects of limitations in the evidence. Few data are completely valid and reliable. The amount of attention to systematic exploration must follow a cost-benefit analysis. Where is the payoff from increasing objectivity going to be great enough to balance the costs of work involved in getting that objectivity? Conversely, where will the harm be sufficiently great due to inaccuracy to require a greater investment in the evaluative process?

Judgment. Judgment is the part of the evaluation process in which alternatives are weighted, comparisons made, various conclusions considered, and worth or value assigned to what's being evaluated. Judgment is a mental operation involving comparison and discrimination through which values and relationships are formulated. Evidence and criteria help in forming sound judgment, but neither constitute judgment. The quality of evaluation rests primarily with the ability of people to use criteria and evidence in forming sound judgment.

Evaluation is an everyday process used by most people. That process can be improved by increasing the information base and degree of system in the procedures used. However, for the evaluation to have "payoff," it has to be internalized by those involved. Therefore, the challenge is helping a process that must be subjective (within the mind) become more accurate.

Procedures based on this model must be designed specifically for given situations. Generally they include:

1. Identifying the purpose of the evaluation. (If it's to be input into decision making, what are the alternatives that need to be compared?)
2. What crucial judgments must be made to complete the purpose?
3. What criteria are germane to those judgments?
4. What type of evidence is needed? How "pure" must it be?
5. What is the most efficient source of such evidence?
6. How does the evidence compare with the criteria?

7. What are the resultant judgments?
8. What do they mean in terms of the purpose of the evaluation?
9. What will recommendations or decisions based on those judgments mean to those involved?
10. If they're to result in action, how can that best be accomplished?

The natural process can use any of several criteria categories as a source of criteria--for example, appropriateness, contact, effort, effectiveness, overall value. It can be applied to any of the programming processes, activities, or stages. However, greatest value is usually secured from good evaluation of plans and preparation before implementing the program and kind of evaluation during the program that has immediate feedback and chance to improve the program while it's in operation.

Evaluation (the assessment of value) provides important input into three kinds of decisions: decisions inherent in the development of programs, decisions inherent in the improvement of programs, and decisions inherent in the fate of the program and programmer. The amount of attention to program components other than results depends on the kind of status that the program is in. For example, evaluation of context, input, and process are particularly important when the program is in a development or improvement status and there's no question but that it will continue. Then result data are important in improving program operations. However, if the fate of the program and/or programmer is at stake, then evaluation of components other than results are secondary to examining results. Some attention is important however so that result data can be accurately interpreted.

SOURCE

Steele, Sara M. "Program Evaluation--A Broader Definition." *Journal of Extension*, VIII (Summer, 1970), 5-17.

Steele, Sara M. "A Contemporary Concept of Program Evaluation." Madison, Wisconsin: University of Wisconsin, Extension, 1972. [Mimeographed]

4.4 MONITORING EVALUATION

SOURCE	Extension education.
ESSENCE	Monitoring evaluation deals with the <i>critical management tasks</i> in keeping programs on track.
USE	Particularly useful in programs you're doing for the first time or in programs that have fallen into a traditional routine.
STARTING POINT	Start with the program plan, statements of the expected outcomes, and standards for the processes to be used.
DESCRIPTION	Evaluation shouldn't be confined to the end of the program process, but should be used continuously. Monitoring is <i>continuous</i> rather than episodic. Monitoring evaluation is like using a combination of a road map and observations of local landmarks in determining how far you are from your destination. The importance of using systematic monitoring depends on: (1) the program's responsiveness to direction and (2) the degree of risk associated with failure of the program. Which activities are irreversible and therefore need to be examined before they reach the point of no return? Major elements of the program used in the model are: <i>time</i>, <i>program action</i>, and <i>expected transitional outcomes</i>. All programming occurs within <i>time frameworks</i>. Program action occurs over time. Expected transitional outcomes exist in time. <i>Expected transitional outcomes</i> describe what must occur at various time slots so the ultimate outcomes will be attained. They're substeps distributed through time and become a source for criteria and standards for monitoring the program. Evaluation cycles should occur as part of each decision in the programming process. Evaluation cycles may also increase the amount of decision making that's done. An evaluation cycle consists of four stages: 1. <i>Design</i>--deciding on what criteria and standards to use to determine the success or progress to that point, 2. <i>Observations</i>--deciding on what data to get. Data may come from direct viewing or from collected information.

3. *Judgment*--deciding what adjustment, if any, needs to be made in the program.
4. *Adjustment*--using the feedback in adjusting the program.

The rapidity with which evaluation cycles will occur depends on the frequency with which decisions are required, the amount of activity in the programming process, the number of subsets of transitional outcomes, and the degree of anxiety with respect to the outcome.

SOURCE

Bruce, Robert L. "Monitoring Evaluation." Ithaca, New York: Cornell University, 1972. [Mimeographed]

4.5 IMPROVEMENT EVALUATION

ORIGIN	Indiana adult basic education.
ESSENCE	This appraisal procedure was developed at the request of state directors. They selected priority areas for examination where they felt evaluation could <i>improve</i> programs.
USE	Excellent when improvement rather than proof of results is desired.
STARTING POINT	Identify the purpose of the evaluation.

DESCRIPTION

One state's program is seldom better or worse than other states. Each has its own unique problems. When the purpose of evaluation is improvement rather than external justification or an attempt to improve status, the approach can be more relaxed and less trauma-producing at the local level. The judgment of outsiders can play an important catalytic role in helping local programs identify areas where improvements need to be made, identify alternatives in areas where improvements must be made, and reinforce confidence in those areas where the program excels.

The following process is indicated:

1. Those commissioning the evaluation meet and define the *specific purposes* that the evaluation's expected to achieve.
2. Areas to be examined are identified and *priorities set*. The sponsoring group develops a list of aspects that might be examined. Priorities are established by each member indicating how he sees priority; then the various views are combined. In the Indiana example, the areas that were selected were: priority 1: curriculum, reading program, recruitment, teacher training; priority 2: administrative and supervisory control, administrative relationships, community relationships, counseling program, learning laboratories, placement, public relations, teacher use of human resources. Other areas that were suggested, but finally classified as non-priority were: mathematics program, relationship of local to state staff, staff relationships, teacher use of material resources, program objectives, student records, and teacher attitudes.
3. Consultants are selected according to the expertise needed in the highest priority areas.

4. Which local programs will be examined is determined through a sampling procedure. In the Indiana example, part of the sample was purposely stratified with the state staff selecting programs according to rural-urban location and geographic representation. An equal number were chosen by a table of random numbers. More than half of the 34 local programs were included.
5. Consultants use a set of general guidelines but are free to assess and comment on any aspect that impresses them either for its strength or weakness.
6. Observations are made by visiting classes; observing students; talking with teachers; talking with students; contacting community citizens; talking to dropouts; talking to superintendents, counselors, supervisors, ABE directors, and relevant others.
7. Consultants file written reports with case examples and recommendations. The reports are combined into an overall report that briefly discusses and gives recommendations on each of the priority areas and then summarizes the overall strength and areas where improvement is needed.
8. The report becomes source material for state level programmers in developing in-service training and in making practical suggestions to local programs.

The on-site consultant activity can be supplemented by:

1. A study of needs in a rural and urban area. Interviews with potential clientele indicate the strength of need and some of the barriers to reaching the clientele.
2. Use of previous reports and studies so that the new effort builds on what's already known rather than repeating such efforts.
3. Examining objectives of programs and/or alternative types of information. In the Indiana example, an earlier study had indicated that only part of the programs had clearly defined statements of objectives. Rather than asking for such statements, local programs were asked to provide one example of what they considered the best achievement of their program and one example of good individual gain through the program. It was assumed that these data would reveal the objectives of the program.

SOURCE

Kreitlow, Burton W. *Indiana Adult Basic Education--An Improvement Evaluation*. Indianapolis, Indiana: State Department of Public Instruction, 1972.

4.6 TRANSACTIONAL EVALUATION

ORIGIN	Evaluation of local programs.
ESSENCE	Transactional evaluation considers the <i>beliefs</i> and <i>feelings</i> of people as they react to suggested changes.
USE	Particularly useful in helping program personnel accept new challenges. It can be of particular help when you're doing a "transplant operation," that is, introducing a change that was developed outside of the system that will be using it, or when outside groups are pressuring for changes and staff are slow or reluctant to make these changes.
STARTING POINT	Identify how people will react to a suggested change in programming.
DESCRIPTION	Transactional evaluation looks at the effects of changed programs on the changers themselves. Its focus is on evaluating program acceptance. It contributes to program improvement to the extent that it facilitates needed changes. Transactional evaluation considers the beliefs and feelings of people. It requires that both the supporters and the critics of a program--those who will be favorably and those that will be adversely affected--be included in the planning and carrying out of the evaluation, and in the forming of conclusions. It's concerned not only with the improvement of a given program, but with the program's effect on the system. In particular, it's concerned with dysfunctions that may occur. Attention is given to examining: 1. Unanticipated as well as expected consequences. 2. Effect on the total organization and not just those directly involved. 3. Dislocation due to competition for resources (including student time and attention) and shifts in roles and expectations. This examination is seen as a continuous process in which the effects of the program are examined by both supporters and critics.

Transactional evaluation differs from traditional summative and formative evaluations in that:

1. The target of evaluation is different. The subject of the evaluation is the system, not the client or the services rendered.
2. The variables relate to the social, psychological, and communications aspects of the system more than to the manifest objectives.
3. Information would be continuously fed back into the system.
4. The evaluator would be more a part of the operating system.
5. Conventional considerations of reliability, validity, and objectivity would be less important than timeliness, relevance, and observable effects of the generation of evaluation information.
6. The aim wouldn't be that of production of new knowledge or the attribution of causality, but would be that of transforming the conflict energy associated with change into productive activity, and the clarification of the roles of all persons involved in changes in the program.

SOURCES

Rippey, Robert M. "Can Evaluation Be Integral to Innovation?" In *Curriculum Evaluation: Potentiality and Reality*, Joel Weiss, ed. Curriculum Theory Network Monograph Supplement. Ontario, Canada: Ontario Institute for Studies in Education, 1972, pp. 45-58.

Rippey, Robert M. "Introduction: What Is Transactional Evaluation?" Paper presented at the meeting of the American Educational Research Association in Chicago, 1972.

GROUP 5

RESULTS--ATTAINMENT OF OBJECTIVES

OVERVIEW:

Evaluation approaches that deal with examining results as specified in objectives are the most common and well developed of all evaluation models. Rather than including individual models in this section, approaches are grouped to indicate areas where development is going on. Distinction is made between approaches that are primarily concerned with the results of instruction (specific behavioral change) and those that focus on programs.

EXAMPLES INCLUDED:

- The *Tyler model* and its adaptations.
- *Instructional evaluation approaches.*
- *Program objectives approaches.*
- *Objectives: categorizations and criteria.*

GENERAL ADVANTAGES:

- Using objectives to provide the major emphasis for examining results makes the whole evaluation process much more efficient.
- Putting some degree of emphasis on objectives also helps the program unit increase its skill in identifying what objectives are reasonable in what programming situations and develop its skill in using objectives as important tools in programming.
- The question probably isn't one of whether objectives should be used in examining the results of programs. It probably should be whether any results other than those specified in the objectives should be examined.

GENERAL CAUTIONS:

- Models that limit themselves to examining the extent to which program objectives have been attained often miss evidence of other important results. Unexpected side effects can be either positive or negative, but often considerably enhance or detract from the overall success of the program.

- All models that invest evaluation resources in relation to objectives assume (1) that the objectives themselves are well developed, realistic, and attainable; best suited to existing needs; and the best choice among various objectives that could be attained; and (2) the design and implementation of the program is such that sufficient input and interaction occurred for the objectives to have actually been attained. Too often, however, one or both of those assumptions fail to be correct and any elaborate data-gathering procedure that focuses only on objectives produces little that's of value to the program.
- Evaluation procedures that stress the whole group attaining the same objective may be inconsistent with the philosophy of education that places the emphasis on individual needs and individual differences. In the latter case, the objectives used should be tailor-made to and probably by the individual student. Then judgments of programs would be on the extent to which a program helped an individual achieve his objectives rather than only on the extent to which a group of people were brought to attain an objective set for the total group.
- Other limitations come not so much in the models themselves but in the intent in which they are used. Models often delete or de-emphasize the steps in the original Tyler model that relate to analyzing the strengths and weaknesses of the program by using data about the attainment of objectives.

5.1 TYLERIAN MODELS

ORIGIN	General education.	
ESSENCE	Evaluation is concerned with determining whether education is actually producing the results that it sets out to achieve as indicated by statements of objectives.	
USE	Can be used in any program that has objectives. However, it's most apt to be successful when there's evidence that the objectives are realistic and the program input great enough to make it logical to expect results.	
STARTING POINT	Analyze the objectives to determine exactly what results are expected.	
DESCRIPTION	<i>THE ORIGINAL TYLER MODEL</i> The role of evaluation is seen as improving curriculum by finding out how far the learning experiences as developed or organized are actually producing the desired results; identifying strengths and weaknesses of teaching plans.	The role of evaluation is seen as improving curriculum by finding out how far the learning experiences as developed or organized are actually producing the desired results; identifying strengths and weaknesses of teaching plans.

The evaluation process includes:

1. Analyzing objectives to identify and clarify their two basic dimensions. Those dimensions are:
 - (a) the behavior expected from the student and
 - (b) the key content that is to be mastered.
2. Identifying situations that will give the student a chance to express the behavior related to the content.
3. Selecting or developing instruments that will record the behaviors.
4. Analyzing the amount of change that's taking place in students.
5. Analyzing the strengths and weaknesses of the program in helping students achieve the objectives.

Prerequisite to this process, objectives have been screened against the needs indicated by various sources--the student, society, and the discipline--and the screens of philosophy and learning theory to be sure that they're appropriate. The learning experiences included have also been evaluated against criteria. This type of evaluation, then, focuses on the behavior of students.

Evaluation must appraise the behavior of students. This appraisal must occur more than once to see whether a change has occurred. Any valid evidence about the specific behavior provides an appropriate method of evaluation. Sampling is important in many ways including a sample of the student's work. Several pieces of evidence will be used to form profiles rather than relying on one piece of evidence.

Curriculum planning is a continuous process. As materials and procedures are developed, they're tried out, results appraised, inadequacies identified, and suggested improvements indicated. Replanning, redevelopment, and reappraisal are a continuing cycle.

SOURCE: Tyler, Ralph W. *Basic Principles of Curriculum and Instruction*. Chicago: University of Chicago Press, 1950.

NATIONAL

ASSESSMENT PROGRAM

Although developed for specific curriculum, the Tyler framework has recently been used for a national assessment. The plan for the National Assessment Program developed by the Exploratory Committee on Assessing the Progress of Education made the following decisions that carried out the basic model:

1. Focus would be on age levels. Four were chosen.
2. There would be no attempt to evaluate individuals or specific school systems or programs. Sample would be of the whole age population with subsampling stratification based on sex, region of the country, type of community, and socioeconomic background, but not school systems.
3. Educational objectives determined what data as to educational results would be collected. Objectives selected were those that: (a) school systems said they were reaching, (b) scholars considered authentic in the various disciplines, and (c) discerning lay adults thought were significant. Considerable negotiation was required to get agreement within and among the three groups.
4. For each objective there was an attempt to describe: (a) things that almost all persons of the age level have accomplished, (b) things that average persons have achieved, and (c) things the most advanced at that age level could do.
5. The data-gathering exercises were packaged so that different students did different things. For example,

30 students spending 40 minutes taking different sets of exercises would provide 1 response to each of a total set of exercises that would take 1 student 20 hours to complete. With 15,000 students responding to the various sets, there would be 500 responses for each set of exercises.

SOURCE: Merwin, Jack C., and Frank B. Womer. "Evaluation in Assessing the Progress of Education To Provide Bases of Public Understanding and Public Policy." In *Educational Evaluation: New Roles, New Means*, Ralph Tyler, ed. Chicago: The University of Chicago, 1970, pp. 306-12.

*FOUR-QUESTION
APPROACH*

Gottman and Clasen define evaluation as *quality control* of the processes and outcomes of an educational program.

They built an evaluation text around four Tyler questions:

1. *Why?*--What needs can you cite that justify the existence of this educational program?
2. *What?*--What are your objectives in the program? That is, what objectives will the program accomplish to meet the need under *why*?
3. *How?*--How will you have the program function to meet its objectives?
4. *How Will You Know?*--What kinds of information should be gathered so that you know if the *how* is meeting the *what* for the *why*?

They suggest the following activities in relation to the questions:

<i>Why?</i>	Needs assessment.
<i>What?</i>	Writing measurable objectives and designing measurement procedures.
<i>How?</i>	Flowchart.
<i>How Will You Know?</i>	Design and use quality control procedures. Emphasis: time-series analysis procedures.

SOURCE: Gottman, John Mordechai, and Robert Earl Clasen. *Evaluation in Education: A Practitioner's Guide*. Itasca, Illinois: F. E. Peacock Publishers, Inc., 1972.

ADOPTION OF THE
TYLER MODEL BY
ADULT EDUCATION

During the 1950s and 1960s, adult education including Extension, pretty well adopted the approach presented in the Tyler model.

Although it usually was not called program evaluation, it served in that capacity.

In 1952, the Committee on Evaluation of the Adult Education Association formulated this definition: "Evaluation is the process of assessing the degree to which one is achieving his objectives. It is looking at one's present position in regard to one's goal. It is a comparison of the actual with the ideal." The committee listed the following principles:

1. Self-appraisal usually is better than appraisal by outsiders.
2. Comparison of achievement with objectives within a program leads to more growth than comparison of one program with others.
3. Everyone concerned with the educative process should be involved in evaluation.
4. Evaluation offers greatest potential benefit if it's a long-time, continuous, and built-in part of the total educational process.
5. Evaluation should be more concerned with results than with energy spent.
6. Objectives should recede as they're approached.

In the chapter on evaluation in a book on outlines of an emerging field of university study, Theide defined evaluation as the process of determining the extent to which educational objectives have been attained. He listed five steps in the evaluating process: (1) determining what to evaluate, (2) defining the behavior desired, (3) determining acceptable evidence, (4) collecting evidence, and (5) summarizing and evaluating the evidence.

Objectives were seen as being essential and derived from values and needs of society, purposes of institutions, needs and interests of individuals, the subject matter itself, and learning theory. Objectives should be achievable, should be in harmony with other objectives of the educator, should be able to be built on for further objectives, should be agreed on and have common meaning, and should be closely related to desired learner behavior. Objectives need to be expressed in specific behaviors desired in the learner.

Situations must be contrived that will permit the learner to exhibit the desired behavior. These situations must be appropriate and relevant, must control unwanted factors, and must be practical and reliable.

Records of behavior can be made by the learner himself, by another observer, or by mechanical devices and should be objective, reliable, and valid. When they've been summarized by counting, describing, and analyzing, judgments are made to show the extent to which objectives have been attained, appropriateness of the objectives themselves, and the effectiveness of the learning experiences. Evaluation must always suit the purposes to be achieved and it must be practical.

The adult learner is task oriented, has specific objectives of his own in mind, and is less willing to accept unknown objectives or the objectives of others. He, therefore, needs to be involved in the establishing and accepting of objectives in the learning situation. Self-evaluation needs to be developed and experimented with in adult education and incorporated in the learning process. Variants in adult education programs create special problems. The tendency for programs to be informal, of short duration, and based heavily on learner objectives makes evaluation by outsiders less acceptable and makes measurement itself more difficult because of the coarseness of available devices and techniques.

Miller and McGuire used the Tyler model as a basis for studying methods for evaluating liberal adult education. They used three criteria for the instruments that they developed.

- Realism:*** The closer we can come to the realities of adult life in our testing, the better able we will be to claim that we've changed students' habitual cognitive behavior. But more than that, tests that are clearly related to real-life problems will also be more interesting and reduce resistance to the testing situation.
- Interest:*** Every possible way of making the instruments themselves interesting to work on will increase the chances of their being used. They should challenge and involve the person who takes them.
- Flexibility:*** The several evaluation devices should be useful to the widest variety of educational programs, and therefore need to be developed in such a way that any individual program may select from them those parts that are particularly relevant.

A manual produced by the Extension Research and Training Division for Cooperative Extension Employees guided Extension in combining a Tyler approach with research-type efforts for examining results.

In 1966, Sabrosky wrote the chapter on evaluation for the new book on the Cooperative Extension Service. She quoted the Tyler definition of evaluation, but discussed briefly the evaluation of organization objectives and evaluation of personnel before examining the place of evaluation in Extension program development. She outlined a series of questions to be examined in evaluation during the planning stage and a parallel set to be examined in terms of the resultant program:

In *planning* for Extension education, the situation is evaluated. In doing so, these questions are asked:

1. Who should be helped to move toward an objective?
 - a. What are these people like?
 - b. What problems do they have?
 - c. Which information and skills do they already have?
2. Which problems should these people get help in solving?
3. Which objectives will be the best ones at which to aim?
4. What resources do they have?
5. What are their motivations?

These questions are then asked:

1. How should these people be helped to move toward each objective?
 - a. Which methods will be most appropriate?
 - b. Which subject matter will be most appropriate?
 - c. When should the teaching be carried out?
 - d. Who should do the teaching?

In *evaluating* Extension education, processes, and results, these questions are asked:

1. Who were helped to move toward an objective?
 - a. What were they like?
 - b. What problems did they have?
 - c. What did they learn? (Behavioral changes)
2. Did these behavior changes seem to help improve the situation?
3. What resources did they lack?
4. What were their motivations?

These questions were then asked:

1. How were the people helped to move toward an objective?
 - a. Which teaching methods were most effective? Why?
With whom? When? By whom? Least effective?
Why? With whom? When? By whom?
 - b. Which subject matter was most useful? Why?
With whom? When? By whom?
 - c. What changes are implied for future Extension programs?

In 1970, Boyle and Jahns defined program evaluation as the determination of the extent to which the desired objectives have been attained or the amount of movement that has been made in the desired direction. They saw it as being primarily concerned with assessing the ends that were attained and not directly concerned with assessments about other aspects of the program such as the means used to attain objectives. Seen in this light, effective evaluation requires: (1) clear, concise objectives or statements of intended, educational ends to be attained; (2) bench mark or pre-program measures of the behavior(s) or behavioral patterns of the learner before his exposure to the educational program; and (3) measures after completion of the educational program.

Sanders (1972) indicates that program evaluation could be conducted for any one or a combination of the following purposes:

- To determine the situation.
- To determine the quality of performance.
- To determine the effectiveness of learning experiences.
- To determine the extent to which objectives are attained.

The existing situation is compared with what the situation should be. What was done is compared with what should have been done. Effectiveness of learning experiences is examined by comparing actual influence with potential influence or one experience with another for relative effectiveness. In determining the extent to which objectives are attained, the behavior expected is compared with the behavior accomplished.

Among other recent adult education texts, Houle (1972) holds closely to evaluation being the determination of the extent to which the educational objective has been accomplished. Axford (1969) focuses on three principles and examples of how to operationalize them. They are: (1) instruments for self-evaluation are preferred to evaluation by others, (2) involving program planners in the evaluation process is necessary, and (3) evaluation should be concerned with outcomes and results rather than activity or energy spent. Klevins chose only to include a chapter of self-evaluation which deals primarily with teacher performance in his 1972 book.

Cass (1971) broke from the prevailing pattern and suggests that evaluation of adult basic education should start with a study of the target population and the number of students enrolled. Other means of evaluating programs include: number of dropouts, number completing each achievement level, number going on to a higher level, persistence of attendance, estimation by teacher and guidance person of student's interest, number enrolling in other schools, follow-up on employment records. Evaluation should be continuous and include information about how adequately each student is attaining his goal(s), which methods are most effective, what content is most satisfying; which materials are most useful, what else teachers need to know.

There's evidence that adult and extension education is using newer models developed in education: for example, Wedemeyer, 1969; Raudabaugh, 1970; Farmer *et al.*, 1970; Hale, 1970; Elliott, 1972. And, as work in this monograph attests, some adult educators are striking out and developing models that are better adapted to adult education. Many of these models are coming from experiences in evaluating programs for disadvantaged adults and evaluating other specially funded federal programs.

Knowles (1970) suggests that evaluation is a much over-emphasized sacred cow and that there's a conflict in values going on. At one end of the spectrum is the emphasis on "hard data" and values of such groups as experimental researchers and cost accountants. At the other end of the spectrum is an emphasis on self-actualization, artistic intuition, free play of natural forces, and creative ambiguity. An adult educator will choose the kind of evaluation that fits his philosophy.

As adult and extension educators move beyond the concept of program evaluation held in the 1960s, some may want to return to the older literature of adult education. For example, Kempfer suggests a framework for examining the program in terms of its comprehensiveness in comparison with the needs of the community and suggest such questions as: What difference does the program make in the life of the community? Are there fewer broken homes? Less discrimination? Are adults becoming more competent in solving both their individual and their group problems? Essert suggests criteria for desirable self-growth as appropriate in evaluating the effectiveness of adult education programs. Categories include: experience of occupational achievement, experience of understanding, experience of self-government, experience of close fellowship, experience of intermittent solitude.

SOURCES: Committee on Evaluation. *Program Evaluation in Adult Education*. Washington, D. C.: Adult Education Association of the U.S.A., 1952.

Thiede, Wilson. "Evaluation and Adult Education." In *Adult Education: Outlines of an Emerging Field of University Study*, Gale Jensen, A. A. Liveright, and Wilbur Hallenbeck, eds. Washington, D.C.: Adult Education Association of the U.S.A., 1964.

Miller, Harry L., and Christine H. McGuire. *Evaluating Liberal Adult Education*. Chicago: Center for the Study of Liberal Education for Adults, 1961.

Byrn, Darcie, ed. *Evaluation in Extension*. Topeka, Kansas: H. M. Ives & Sons, Inc., 1959.

Sabrosky, Laurel. "Evaluation." In *The Cooperative Extension Service*, H. C. Saunders et al. Englewood Cliffs, New Jersey: Prentice-Hall, Inc. 1966, pp. 339-51.

Boyle, Patrick G., and Irwin R. Jahns. "Program Development and Evaluation." In *Handbook of Adult Education*, Robert M. Smith, George F. Aker, and J. R. Kidd, eds. New York: The Macmillan Company, 1970.

Sanders, H. C. *Instruction in the Cooperative Extension Service*. Baton Rouge, Louisiana: Louisiana State University and Agricultural and Mechanical College, Publishers, 1972, pp. 76-87.

Houle, Cyril O. *The Design of Education*. San Francisco: Jossey-Bass, Inc., 1972, p. 231.

Axford, Roger W. *Adult Education: The Open Door*.
Scranton, Pennsylvania: International Textbook
Company, 1969.

Seaman, Don F., and Donnie Dutton. "Self Evaluation."
In *Materials and Methods in Adult Education*,
Chester Klevins, ed. Canoga Park, California: Klevens
Publications, Inc., 1972, pp. 265-74.

Cass, Angelica W. *Basic Education for Adults*.
New York: Association Press, 1971, pp. 153-57.

Wedemeyer, Charles A. "Evaluation of Continuing
Education Programs." *American Journal of Pharmaceutical
Education*, XXXIII (No. 5), 793-863.

Raudabaugh, J. Neil. "Evaluation As a Concept."
Presentation to the ECOP-EMIS Advisory Committee
Meeting, Lake Texoma, Oklahoma, November, 1970.

Farmer, James A., Jr., Ralph K. Sylvester, and Patrick
J. Weagraff. *Western Region AMIDS Evaluation*.
Los Angeles: University of California, Division of
Vocational Education, 1970.

Elliott, Elizabeth. "A Model for Evaluating Educational
Programs Aimed at Disadvantaged Families." Paper
presented at the Adult Education Research Conference,
Chicago, Illinois, April, 1972.

Hale, Larry A. "Evaluation Is a Process of Deciding the
Value of Something." Paper prepared for the Missouri
Extension Program Development Conference, 1970.

Knowles, Malcolm S. *The Modern Practice of Adult
Education*. New York: Association Press, 1970.

Kempfer, Homer. *Adult Education*. New York: McGraw-Hill
Book Company, Inc., 1955, pp. 404-19.

Essert, Paul L. *Creative Leadership of Adult Education*.
Englewood Cliffs, New Jersey: Prentice-Hall, Inc.,
1951, pp. 261-87.

5.2 INSTRUCTIONAL EVALUATION APPROACHES

ORIGIN	General education.
ESSENCE	Precision in stating objectives facilitates instruction and evaluation.
USE	Particularly useful in evaluating changes in knowledge.
STARTING POINT	Specify objectives in very precise terms.

DESCRIPTION

Much of the current work on objectives and many of the recent interpretations of "educational" evaluation are specific to the evaluation of instruction rather than sufficiently broad as to encompass program evaluation.

However, examining some of the more recent developments in instructional evaluation is useful in that some programs are made up primarily of instructional type activities (that is, classes or courses) and some of the basic ideas in instructional evaluation, for example, setting standards of performance for objectives can be adapted to program evaluation.

GOAL-REFERENCED INSTRUCTION

Popham and Baker have succeeded in translating a fairly complex curriculum procedure into a model acceptable to the average teacher. They suggest that a goal-referenced instructional model has four main parts:

1. Specification of objectives.
2. Preassessment.
3. Instruction.
4. Evaluation.

Objectives are to be stated in measurable learner behavior. Two kinds of techniques can be used to judge the degree to which the learner has attained the instructional objective: (1) observation of behavior and (2) examination of products that the learner will produce. For example, observing the student giving a speech (behavior) or examining an essay or test that he has written (product).

There are two dimensions to setting levels of performance in regard to an objective. One is the level of performance

for each student, the other is the level of performance for a group of students. Examples of such standards are: individual student, each student should attain scores of at least 90; group of students, 90 percent of the class should attain a score of 90 or better. Standards should be set for the minimal acceptable attainment.

Objectives need to be sequenced for effective instruction. This sequencing helps in later decisions about how attainment will be evaluated. The following guides have been suggested:

1. Analyze objectives into content-behavior modules, not just content.
2. Use guidance from a regularized scheme such as that advocated by Gagné or Bloom.
3. To arrive at a first sequence, repeatedly ask the question: "What does the learner need to be able to do before he can perform this task?"
4. Clearly identify prerequisites not within your instructional responsibility.
5. Where possible, empirically verify sequences generated.

After tentative objectives have been set, the teacher should use preassessment procedures to determine how many of those objectives are presently attained by the majority of the students and where further emphasis is needed by the majority and remedial work by a minority. Preassessment helps prevent reteaching when such teaching isn't necessary and helps identify where present emphasis should be placed. It also tests assumptions about the kinds of prerequisites students will be bringing to the learning situation.

The preassessment helps identify whether differentiated instruction is necessary. Such instruction provides two or more tracts suited to the general abilities within the group. Two of the most common ways of differentiating are: (1) setting additional objectives for more advanced students or (2) identifying more challenging means for accomplishing the same set of objectives.

Evaluation at the end of the teaching sequence consists of securing evidence either of the behavior of students or of examining products of that behavior. This evidence can be secured either under natural conditions or under manipulated conditions. Listening to grammar used at break time would be an example of securing evidence under natural conditions.

Tests are usually conducted under manipulated conditions in that the teacher deliberately sets up the stimulus condition. Tests are of two kinds. Some are norm referenced. Others are criterion referenced. Norm-referenced tests are those where individual scores are interpreted in relationship to the scores of others. Criterion-referenced tests compare performance to a set criterion without regard to the scores of others. Criterion-referenced evaluation of student performance differs from the usual grading discrimination system. A criterion-referenced approach to examining the performance of adult students may be more appropriate than is the traditional procedure of comparison with the work of others.

In a guidebook developed under contract to the Bureau of Indian Affairs, Popham defines evaluation as the act of *assessing merit* by judgmentally comparing (1) the observed result (performance data) of some educational enterprise with (2) a desired standard or criterion of acceptability (preference data).

He suggests two major roles for this type of evaluation:

Role One--Educational Needs Assessment. Determining the desired *ends* of the educational system.

Role Two--Treatment Adequacy Assessment. Judging the worth of educational *means*.

Formative. To improve the treatment as it's being developed.

Summative. To compare the treatment with competitors.

The following guidelines for educational evaluators are presented after a discussion of the key ideas involved in the guideline:

1. The educational evaluator should encourage the use of instructional objectives that provide explicit descriptions of the post-instructional behavior of learners.
2. While recognizing that nonmeasurable goals will be of limited use for his purposes, the educational evaluator must be aware that instructors may wish to devote a reasonable proportion of their efforts to the pursuits of important but currently unassessable objectives.
3. The educational evaluator must identify criteria of adequacy when using instructional objectives that require constructed responses from learners.

4. The educational evaluator should foster the use of measurable objectives that possess content generality rather than test-item equivalence.
5. Before the introduction of the instructional treatment, educational evaluators should strive to establish minimal proficiency levels for instructional objectives.
6. The educational evaluator will often find the taxonomies of educational objectives useful both in describing instructional objectives under consideration and in generating new objectives.
7. The educational evaluator should consider the possibility of selecting measurable objectives from extant collections of such objectives.
8. The educational evaluator should avoid the use of norm-referenced measures, preferring instead criterion-referenced measuring devices.
9. The educational evaluator should use domain-referenced achievement testing procedures when called on to generate test items for objectives or to determine the content validity of already developed test items.
10. The educational evaluator should encourage the use of multiple criterion measures by generating alternative schemes for assessing learner behaviors.
11. When the use of customary measuring tactics may produce reactive effects, the educational evaluator should employ unobtrusive measures.
12. The educational evaluator should be particularly attentive to unanticipated outcomes that result from an instructional treatment, and should use such outcomes in judging the merits of the treatment.
13. The educational evaluator should clarify value preferences of various groups regarding desired educational goals by having sets of precisely stated objectives rated by the individuals involved, then translating these ratings into composite indicators of each objective's worth.
14. The educational evaluator can compare preference and performance data by contrasting preference ratings of objectives with measures designed to assess the degree to which learners have achieved those objectives.

15. Whenever large scale measurement is required, the educational evaluator should consider the economic advantage of employing item-sampling and person-sampling procedures.
16. In conducting formative evaluations, the educational evaluator will find it useful to (a) employ small samples of learners, (b) secure performance data about terminal and en route objectives, and (c) use designs such as the one-shot case study or the one group pre-test/post-test design.
17. In conducting summative evaluations, the educational evaluator should prefer the use of designs involving randomized control groups, that is, the pre-test/post-test control group design or the post-test only control group design, but when randomization is impossible can profitably employ the nonequivalent control group design or the interrupted time series design.
18. The educational evaluator should analyze data according to the smallest independent units available, frequently leading to the use of classroom or larger units rather than individual pupil units.
19. The educational evaluator should, in general, prefer descriptive statistics and estimation procedures instead of statistical hypothesis testing procedures.
20. The educational evaluator should present decision makers with a wide range of pertinent information so that choices among alternatives can be made in a cost-effectiveness context.

SOURCES: Popham, W. James, and Eva L. Baker. *Systematic Instructions*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1971.

Popham, W. James. *An Evaluation Guidebook*. Los Angeles: The Instructional Objectives Exchange, 1972.

**EVALUATION AS
FACILITATION OF
LEARNING**

A 1971 handbook by Bloom, Hastings, and Madaus is concerned with putting the evaluation of instruction and the evaluation of learning back together

and improving learning as a result. The handbook is viewed as encompassing:

1. Evaluation as a method of acquiring and processing the evidence needed to improve the student's learning and the teaching.

2. Evaluation as including a great variety of evidence beyond the usual final paper and pencil examination.
3. Evaluation as an aid in clarifying the significant goals and objectives of education and as a process for determining the extent to which students are developing in these desired ways.
4. Evaluation as a tool in education practice for ascertaining whether alternative procedures are equally effective in achieving a set of educational ends.

The authors present a two-dimensional figure showing the relationships among analysis of the learners, instructional decisions and evaluation (row headings); and student, instructional process, and objectives (column headings).

The three stages of evaluation that correspond to the column headings and some of their activities are:

Initial Evaluation

1. Record of previous relevant achievement.
2. Diagnostic and placement tests.
3. Aptitude tests relevant to the choice of alternative learning approaches.

Formative Evaluation

1. Tests relevant to the structure of each learning unit or task.
2. Analysis and diagnosis of what must still be learned.
3. Prescription of alternative learning materials and approaches needed to master the unit or tasks.

Summative Evaluation

1. A sample of the attainment of objectives and content in the outcome model.
2. Analysis and diagnosis of extent to which various parts of the outcome model have been attained.
3. Analysis and diagnosis of extent to which each learner has attained the outcomes on the table of specifications.

The outcome model or table of specifications is a grid that lists the key elements of content of the course or unit on one dimension and the expected behaviors related to that content on the other dimension. Cells are then formed coupling behaviors with content to indicate potential objectives. The cells actually used to frame the expected outcomes are judged in terms of whether they are both possible and desirable.

SOURCE: Bloom, Benjamin S., J. Thomas Hastings, and George F. Madaus. *Handbook on Formative and Summative Evaluation of Student Learning*. New York: McGraw-Hill Book Company, 1971.

ADDITIONAL REFERENCES

Blaney, John P. "The Application of Formative and Summative Evaluation Techniques to Short Term Adult Education Courses." Paper presented at the Adult Education Research Conference, Chicago, Illinois, 1972.

Popham, W. James, Elliot W. Eisner, Howard J. Sullivan, and Louise L. Tyler. *Instructional Objectives*. AERA Monograph Series on Curriculum Evaluation, No. 3. Chicago: Rand McNally & Company, 1969.

Evaluation Comment. Los Angeles: University of California, Center for the Study of Evaluation. [Periodical]

Wilhelms, Fred T., ed. *Evaluation As Feedback and Guide*. Washington, D.C.: Association for Supervision and Curriculum Development, NEA, 1967.

Wittrock, M. and D. Riley. *Evaluation of Instruction: Issues and Problems*. New York: Holt, Rinehart and Winston, 1970.

5.3 PROGRAM OBJECTIVES APPROACHES

ORIGIN	Education and other social programs.
ESSENCE	This group of approaches combines objectives with other elements such as criteria or program components.
USE	Helpful in getting a broader understanding of how programs achieve objectives.
STARTING POINT	Define the objectives and identify other relevant components.
DESCRIPTION	This set of evaluation frameworks either specifically says it's concerned with objectives that go beyond the kind of objectives that are found within a specific course or instructional unit, or is so phrased that it can be applied in a broader setting. <i>CRITERIA OF SUCCESS APPROACH</i> "Evaluation is the determination (whether based on opinions, records, subjective or objective data) of the results (whether desirable or undesirable; transient or permanent; immediate or delayed) attained by some activity (whether a program or part of a program . . . an ongoing or one-shot approach) designed to accomplish some valued goal or objective (whether ultimate, intermediate, or immediate, effort or performance, long or short range). This definition contains four key dimensions: (1) process--the 'determination,' (2) criteria--'the results,' (3) stimulus--the 'activity,' and (4) value--the 'objective.' The scientific method with its accompanying research techniques then provides the most promising means for 'determining' the relationship of the 'stimulus' to the 'objective' in measureable 'criteria'." Five general categories of criteria are suggested against which the success or failure of a program may be evaluated. Attention is given to a sequence of objectives from immediate to intermediate to long range. An understanding of the assumptions that underlie the objectives and a checking of the validity of those assumptions is essential. Five general categories of criteria according to which the success or failure of a program in attaining its immediate-, intermediate-, or long-range objectives may be evaluated are: <i>effort, performance, adequacy of performance, efficiency, and process.</i>

Effort. This category includes criteria concerned with the quantity and quality of activity that takes place. It deals with questions like: What was done? How well was it done? This criterion area can be compared with measuring the number of times a bird flaps its wings without trying to determine how far it flies.

Performance. Criteria in this category are concerned with the result of effort. Questions include: How much is accomplished in terms of the immediate goals? Did any change occur? Was the change the intended one? Using this type of criteria is comparable to examining how far the bird has flown instead of merely counting the flappings of its wings.

Adequacy of Performance. This category of success criteria deals with degree to which the effective performance is adequate to the total need. How many of the people who are potential users of the program actually are reached and actually use it? How much impact has the program had on the needs and problems that initiated it? Judgments based on criteria in this category must also be conscious of reality in terms of what can be expected with a given level of knowledge and available resources. In terms of the analogy of the bird, this category is concerned with how far the bird has flown compared to how far it needs to go.

Efficiency. Efficiency is concerned with the evaluation of alternative paths or methods in terms of costs--money, time, personnel, and public convenience. It represents a ratio between effort and performance and deals with such questions as: Is there any better way to attain the same results? In the bird analogy, comparable questions would be: Could the bird have arrived at his destination more efficiently by some other means than flying the way he did? Did he take advantage of air currents? Did he fly too high or not high enough?

Process. The process dimension isn't parallel with the other four categories, but is added because it helps make sense out of evaluation findings. It helps locate weaknesses and identify whether the program can be improved or whether it should be dropped.

Four dimensions of process are suggested: (1) the *attributes* of the program itself, (2) the *population* exposed to the program, (3) the *situational context* in which the program takes place, and (4) the different kinds of *effects* produced by the program. Within the attributes of the program, we search the component parts of the program for the one or ones that most contribute to

or detract from success. The population dimension explores such questions as: Which people are most affected by the program? Who do you succeed in reaching and who don't you reach? Who makes the best target population for a program--the individual, the group, the public? Do you reach them as ultimate target? The third dimension explores whether the conditions greatly affect the success of the program. Could the same program be set up under different conditions and be as successful?

The fourth category, effects, can be broken down in various ways: unitary or multiple effects, unintentional or side effects, duration of effects, and type of effect. In each case, being able to specify the components existing in the actual program situation helps you better understand the success or failure of the program.

SOURCE: Suchman, Edward A. *Evaluative Research*.
New York: Russell Sage Foundation, 1967.

O-A-R
MODEL

Program performance is evaluated by examining the effectiveness and efficiency of the program. Evaluation of both characteristics focus on the relationship of three key components: *objectives*, *activities*, and *resources*. These three components are defined as follows.

Objectives. A situation or condition of people or of the environment that responsible program personnel consider desirable to attain. (Objectives themselves include ultimate objectives, program objectives, and subobjectives.)

Activity. Work performed by program personnel and equipment in the service of an objective.

Resources. Personnel, funds, materials, and facilities available to support the performance of activity. The following abbreviations are used in the model:

- AO Attainment of objectives that can be attributed to the program activity.
- PO Proposed objectives for attainment through the program activity.
- AA Actual activities performed.
- PA Planned activities performed.
- AR Actual resource expenditure.
- PR Planned resource expenditure.

Effectiveness. In general, questions concerning effectiveness are directed toward assessing the extent to which a planned or intended objective has been attained as a result of program activity. An analysis is thus suggested in which the proportion of attainment of program objectives that's attributable to program activity (AO) is compared with the desired level, which during the planning process, the planners had proposed would result from the program activity (PO). Program effectiveness is denoted as the ratio, AO:PO. However, two subordinate measures must be considered to determine the soundness of the assumptions. One is the extent to which the activity has been carried out and has used resources in the way planned. In other words, the ratio of actual activities to planned activities, AA:PA. The second is the ratio of the actual expenditure of resources to the planned expenditure, AR:PR.

Efficiency. Efficiency is the ratio between an output and an input, AO:AR. The reverse, AR:AO, yields a measure of average cost. These measures of overall efficiency may be interpreted by examining two intermediate efficiency measures: activities performed to resources expended, AA:AR or AR:AA; and objectives attained to activities performed, AO:AA or AA:AO.

Continuous evaluation measures can be obtained in the following sequence:

1. The extent to which resources are being expended as planned, AR:PR.
2. The extent to which activities are being performed in the quantity and quality planned, AA:PA and the efficiency of resource expenditures, AA:AR.
3. The net attainment of selected subobjectives, AO_{sub}:PO_{sub} and efficiency of subobjective attainment, O_{sub}:R_{sub} and O_{sub}:A_{sub}.
4. Program effectiveness, AO:PO; program efficiency, AO:AR;; and activity efficiency, AO:AA.

SOURCE: Deniston, O. L., et al. "Evaluation of Program Efficiency." *Public Health Reports*, LXXXIII (No. 7, 1968).

**CRITERIA AND
OBJECTIVES**

This outline for the evaluation process looks as though it should have been included in the section on program models. However, the process is applied primarily to the purpose and objectives of the program.

1. *Purpose of Evaluation.*

- a. Improving a program or specific component of a program.
- b. Proving value and importance of a program.

2. *Criteria Selection.*

The criteria used will depend largely on the program characteristics being evaluated. The program characteristics most frequently evaluated are:

- a. Effectiveness.
- b. Efficiency.
- c. Suitability.
- d. Importance.

3. *Preparation for Evidence Collection.*

- a. Determine program characteristic(s) to be evaluated (mentioned above).
- b. Determine what program level or component is to be evaluated.
- c. Determine program stage(s) to be evaluated.
 - 1) Program determination.
 - 2) Program preparation.
 - 3) Program implementation.
 - 4) Program evaluation.

4. *Evidence Collection.*

- a. Sources of evidence.
- b. Evidence collecting methods and devices.
- c. When to collect the information.
- d. Who should be involved in collecting evidence.

5. *Analysis and Interpretation.*

- a. Organizing and classifying evaluation data.

- b. Making comparisons and judgments.
- 6. *Reporting and Use.*
 - a. Identify audience(s).
 - b. Identify how the report will be used.
 - c. Report format.

Criteria are derived from the objectives. Generally it's more difficult to develop criteria for high level (more general) objectives such as those that describe the purpose of a program than it is to develop them for specific instructional objectives.

Efficiency is discussed in terms of alternative methods for securing the same objective. Suitability deals with the appropriateness of the objectives. Importance deals with the overall value of what's accomplished through attaining these particular objectives.

SOURCE: Matteson, Harold R. *The Process of Evaluating Vocational Educational Programs in Agriculture*. Instructional Series No. 3. Madison, Wisconsin: University of Wisconsin, Department of Agricultural and Extension Education, February, 1972.

LOCALLY DIRECTED EVALUATION

Byram and Robertson present a framework that starts with objectives and gathers evidence about how well they've been attained. Emphasis should be on the program objectives and philosophy of education. Objectives of course and instructional units contribute to but don't constitute program objectives. The next steps deal with formulating criterion questions; identifying and obtaining evidence; analyzing, interpreting, and reporting information; and formulating and implementing recommendations. The process includes examination of input, process, and community needs.

The unique contribution of this particular model is that it recommends that the evaluation involve local people (staff and citizens) in making the evaluation. They list the following essential elements in local evaluation:

1. A commitment to the evaluation effort by the local administration.
2. A strategically placed local leadership team.
3. An active staff steering committee.

4. A functioning citizens' committee, advisory to the staff.
5. An allocation of time to the leadership team.
6. Training in research and evaluation procedures.
7. Communications and visibility.
8. Qualified consultant service.

SOURCE: Byram, Harold M., and Marvin Robertson. *Locally Directed Evaluation of Local Vocational Education Programs*. East Lansing, Michigan: Michigan State University, College of Education, Department of Secondary Education and Curriculum, March, 1970.

OHIO
MODEL

Most attainment of objectives models are applied to specific local programs. However, models have been developed to

apply the approach at a state or national level.

Ohio reviewed existing methodological approaches to the evaluation of vocational education and formulated a system that included:

1. Program objectives and measurable goal statements.
2. A data set, instrumentation, and procedures to measure the effectiveness of efforts to achieve specific program goals.
3. The means by which the evaluation system would be articulated with larger management systems of state divisions of vocational education.

A prerequisite was that the system be designed for self-initiated evaluation that would contribute to decision making involved in state level program planning. A second was that it be a continually operative mechanic so that program plans might be adjusted whenever required by changes in the field situation. The model integrated evaluation and program planning and is called the Evaluation-Program Planning Cycle.

The goal statements for each objective were formulated to permit measurement of the extent to which: (1) target populations of concern are being served and (2) local programs of vocational education assure program comprehensiveness, relevance to desired program outcomes, quality, and assessability.

The data sets were designed to provide quantitative indicators of: (1) target population needs, (2) the effectiveness of training outcomes, (3) school and community characteristics, and (4) the relationships between vocational education program processes and training outcomes.

State level program planners require evaluation results to plan and redirect agency activities and to selectively allocate the personnel and financial resources required to accomplish these activities. The data were organized in ways that would have the greatest usefulness for these management operations. Data were organized according to:

1. Program sectors--public, private.
2. Program levels--secondary, post-secondary, adult.
3. Program areas--trade and industrial, agricultural, etc.
4. Facility types--area vocational school, secondary school.
5. Student characteristics and experiences--equal opportunity data, cooperative work experiences, etc.
6. Schools.

Three efficiency factors received attention during the process of formulating the data set:

1. The time frame required for collecting the data.
2. The costs to the state agency of carrying out data collection procedures.
3. Personnel required to manage the data collections.

SOURCE: Starr, Harold, et al. *A System for State Evaluation of Vocational Education*. Research Series No. 58. Washington, D.C.: U. S. Department of Health, Education and Welfare, Office of Education, Bureau of Research, May, 1970.

**PROGRAM PLANNING
BUDGETING SYSTEM
(PPBS)**

PPBS (Program Planning Budgeting System) is one way that objectives have been used in administration models.

PPBS has four major phases: (1) *planning*, which is concerned with generating program objectives; (2) *programming*,

which is concerned with the preparation of alternative sets of activities and services to achieve program objectives, (3) *budgeting*, which is involved with the formulation of detailed plans, accounting, and reporting; and (4) *evaluating*, which is concerned with progress outputs and effectiveness of programs.

This form of budgeting requires that projected outcomes and evidence of past results be presented along with budget requests. It requires that cost-benefit analysis be done to compare the efficiency of one proposal with another.

At the federal level it involves:

1. Appraisals and comparisons of various government activities in terms of their contributions to national objectives.
2. Determination of how given objectives can be attained with a minimum expenditure of resources.
3. Projection of government activities over an adequate time horizon.
4. Comparison of the relative contribution of private and public activities to national objectives.
5. Revisions of objectives, programs, and budgets in the light of experience and changing circumstances.

The model was introduced at the federal level and has been adopted at state and local levels.

SOURCES: Novick, David, ed. *Program Budgeting*.
Cambridge, Massachusetts: Harvard University
Press, 1965.

Alioto, Robert, and J. A. Jungherr. "Using PPBS to Overcome Taxpayers' Resistance." In *Emerging Patterns of Administrative Accountability*, Lesley H. Browder, Jr., ed. Berkeley, California: McCutchan Publishing Corporation, 1971, pp. 252-59.

SYSTEM VARIABLES APPROACHES

Another kind of modeling relates objectives to the program and/or other factors.

NEA Model

One of the first major approaches to broadening the Tyler model came when Taba and Sawin introduced a model developed by the Association for Supervision and Curriculum Development of the National Education Association.

Taba and Sawin listed the following weaknesses of evaluation as carried on in the schools: (1) the objectives that form the basis for evaluation are usually too narrow, (2) the range of instruments and devices that are being used is often too limited, (3) the focus of attention has been on the end product rather than on process and has resulted in inadequate knowledge of the processes by which the end products are attained, (4) results have often been interpreted without adequate information about factors which affect learning and achievement, and (5) results of evaluation haven't been adequately translated into curriculum decisions.

They proposed four major categories arranged on the four sides of the square. Objectives are on the top of the square with evidence of pupil behavior pertaining to the objectives on the bottom. Evidence of teaching-learning operations and evidence of nonschool factors affecting learning complete the other two sides of the square.

SOURCE: Taba, Hilda, and Enoch Sawin. "A Proposed Model of Evaluation." *Educational Leadership*, XX (October, 1962), 57-59ff.

*Cube
Models*

Hammond's model presents relationships in a three-dimensional cube. One dimension is the *behavioral* dimension with objectives classified into three types: cognitive, affective, and psychomotor. The other two dimensions are the *instructional* dimension (organization, content, method, facilities, and cost) and the *institutional* dimension (student, teacher, administrator, educational specialist, family, and community). The sub-items within the dimensions represent types of variables that may need to be considered in setting objectives and in examining the effectiveness of programs in meeting those objectives.

Armstrong and his colleagues from the EPIC Evaluation Center at Tucson use a similar cube. They say that the institutional, instructional, and behavioral elements all should be considered when objectives are written and their attainment evaluated. In addition, they add the variables of measurement, time needed, and proficiency level.

SOURCES: Hammond, Robert L. *Evaluation at the Local Level*. Columbus, Ohio: Ohio State University Evaluation Center, 1967.

Armstrong, Robert, et al. "A Scheme for Evaluation." In *Educational Accountability Through Evaluation*, E. Wayne Roberson, ed. Englewood Cliffs, New Jersey: Educational Technology Publications, 1971.

Information Domains Nelson suggests that there are three main domains from which information about how a developing curriculum attains its objectives is gathered. They are:

<u>Ingredients</u>	<u>Processes</u>	<u>Product</u>
Learning materials	Teaching acts	Student
Teachers	Learning activities	Teacher
Students	School environment	School
Schools		Parent
Community		

SOURCE: Nelson, Orville. "The Evaluation System for the American Industry Secondary School Courses." Paper presented at the AVERA program at the AVA Convention in Dallas, Texas, December, 1968.

TYPES OF EVALUATION The categorization of types of evaluation as used by the Office of Economic Opportunity is proposed. Those four categories are: program impact evaluation, program strategy evaluation, project evaluation, and project rating.

Program Impact Evaluation. Assessment of the overall effectiveness of a program in meeting its objectives or the relative effectiveness of two or more programs in meeting common objectives. The usual objective of program impact evaluation is to help policy makers in reaching decisions on program funding levels or on possible redirection of the program.

Program impact evaluation depends on the definition and measurement of appropriate output variables and on the use of appropriate comparison groups. Environmental and process data aren't usually essential.

Program Strategy Evaluation. Assessment of the relative effectiveness of different techniques used in a national program. The usual objective of program strategy evaluation is to inform program managers of the relative effectiveness of the different strategies or methods used by projects in the national program.

Program strategy evaluation depends on definition and measurement of appropriate environmental, input, process, and output variables selected on the basis of suitable analytic models.

Project Evaluation. Assessment of the effectiveness of an individual project in achieving its stated objectives. This form of evaluation is required in many federal programs and is often carried out by the project itself. Project evaluation requires measurement of the important output variables as well as the use of comparison groups.

Project Rating. Assessment of the relative effectiveness of different local projects in achieving program objectives. The usual objective of project rating is to provide program managers with information on the relative success of local projects operating within a national program. In most cases, it will make sense to rate projects against one another only if they're operating in similar environment.

Project rating depends on definition and measurement of environmental variables and relatively inexpensive output measures (for example, measure of short-term impact).

Each of the four categories can vary in the amount of detail considered and the level of decision making involved. The relative importance of the four also will vary from program to program and over time, as will the feasibility of conducting such evaluation.

Feasibility should be examined before evaluation is attempted. Whether evaluation is feasible depends on: type of evaluation wanted, availability of theory and measures suitable for the type of evaluation desired, time available for evaluation, availability of suitable comparison groups, feasibility of collecting the required data on program participants and members of appropriate comparison groups. In general, evaluation is more feasible when inputs are tightly controlled or when the effect of the program can be expected to be "large" compared to that of other forces that will be operating.

Monitoring is always feasible from a methodological standpoint. Feasibility of the four types of program evaluation depends on whether suitable measures are available. Project rating is usually feasible. Its main problem will be devising suitable systems for classifying the environments within which projects are operating. Program impact evaluation will usually be methodology feasible.

It's usually impossible or too costly to learn enough from the evaluation of one local project to say with confidence whether the project--or other forces--actually caused the effects measured. For example, success may be due to the charismatic teacher.

Related activities that aren't accepted by the author as being evaluation include:

Monitoring. The assessment of managerial and operational efficiency of programs or projects, through periodic site visits and other management techniques. The usual objective of monitoring is to give program managers impressionistic data about how their projects are going, to see if they're being run efficiently, if they're following program guidelines, if they have competent staff in individual projects.

Reporting Systems. Routine reporting from state or local level isn't evaluation, but may furnish useful data on services provided, populations served, and costs of providing services. In some cases, it may be justifiable to spend evaluation funds to defray state and local costs of providing such data, as well as follow-up data on specified samples of program or project participants, when the data are required to meet clearly specified evaluation needs.

Cost Analysis. Program managers often have good information on the cost (to the federal government) of providing services through a program. Comparative analysis of costs by project, by groups of projects, or by program is essential in good evaluation as well as a valuable management tool in itself.

SOURCE: Wholey, Joseph, et al. *Federal Evaluation Policy: Social Evaluation by Federal Agencies*. Washington, D.C.: The Urban Institute, 1970, pp. 1-9.

5.4 OBJECTIVES: CATEGORIZATIONS AND CRITERIA

ORIGIN	General education.
ESSENCE	Part of the difficulty in evaluating the attainment of objectives is the difficulty in setting and using objectives. Recent efforts recognize the importance of better understanding the phenomena that are objectives.
USE	Particularly useful when people are having trouble setting, using, and evaluating the attainment of objectives.
STARTING POINT	Examine the objective and establish its relationship to other objectives.
DESCRIPTION	Because objectives are the central focus of this approach to program evaluation, the nature of their being must be well understood. Various attempts have been made to describe and categorize kinds of objectives, that is, develop models of types of objectives. <i>OBJECTIVES AS SYSTEMS AND PARTS OF SYSTEMS</i> Objectives may be considered as a system. They have three interrelated parts: (1) the content or the "what" of the objective; (2) the type of behavior or action expected in relation to that content or the "what's he supposed to do with it" component; and (3) the "who" and "how many" components that indicate who's supposed to attain the objective. The three parts interact together and depict the "whole" of the learning that's expected to occur. Each of these three parts hook into other systems. The "who" and "how many" hook into the socio-psychological entities that are learners. The content hooks into systems for analyzing and categorizing content. Gagné, the authors of the taxonomies of educational objectives, and Brunner have been particularly helpful in clarifying how various types of content (concepts, facts, chains of specifics, etc.) relate to each other and constitute a knowledge or attitude system. Furst, Mager, and the taxonomists have also been helpful in seeing how various kinds of behavior fit together into a whole range of behavior. Understanding how the parts of the system making up an objective fit into knowledge, attitude, and behavior systems makes it easier to determine the best way to examine whether the change specified in the objective has occurred.

SOURCES: Furst, Edward J. *Constructing Evaluation Instruments*. New York: Longmans, Green, 1958.

Mager, Robert F. *Goal Analysis*. Felmont, California: Fearon Publishers/Lear Siegler, Inc., Educational Division, 1972.

Mager, Robert F. *Preparing Instructional Objectives*. Felmont, California: Fearon Publishers/Lear Siegler, Inc., Educational Division, 1962.

Krathwohl, D. R., B. S. Bloom, and B. B. Masia. *Taxonomy of Educational Objectives: The Classification of Educational Goals*. Handbook 2: Affective Domain. New York: McKay, 1964.

Bloom, Benjamin S., ed. *Taxonomy of Educational Objectives: The Classification of Educational Goals*. Handbook 1: Cognitive Domain. New York: McKay, 1956.

Gagné, R. M. *The Conditions of Learning*. New York: Holt, Rinehart and Winston, 1956.

Bruner, J. S. *Toward a Theory of Instruction*. Cambridge, Massachusetts: Harvard University Press, 1966.

Although written as an individual unit, an objective seldom stands completely alone. Individual objectives usually are part of some type of network. They usually make up a network in the sense that a set of objectives at the same levels are stated for a particular activity. Although defining separate things, there often is a definite relationship among them. For example, cognitive, psychomotor, and affective objectives are closely interwoven. Adults usually don't bother to exert the effort taken to master content or perfect skills unless they believe in what they are mastering. Sometimes objectives at the same level involve sequence. One has to be accomplished before its possible to accomplish the next. Often a major objective involves a whole package of more specific objectives each of which is essential to the attainment of the larger objective.

The nature of the interrelationships must be understood to adequately evaluate whether objectives have been attained. Two major ways of looking at relationships are longitudinally and hierarchically.

LEVEL AND
TIME SEQUENCE
OF OBJECTIVES

Suchman discusses the relationship of immediate objectives, intermediate objectives, and ultimate objectives.

He points out that chains of objectives exist linking objectives at different levels. The chain includes the idealized objectives or broad mission statements at one end and specific tasks at the other.

Suchman also stresses that assumptions underlying the validity of each level of objectives and assumptions regarding the relationship between levels may need to be rechecked.

Taylor and McGuire model a procedure by which broad objectives become specific student outcomes. They use four stages in their model: (1) broad objectives derived from societal and professional pressures, (2) interpretations into behavioral objectives, (3) translation into forms suitable for the classroom environment, and (4) student outcomes. They view curriculum evaluation as including data of two kinds: (1) objective descriptions of the goals, environments, personnel, methods and content; and immediate- and long-range outcomes; and (2) recorded personal judgments of the quality and appropriateness of goals, inputs, and outcomes.

SOURCES: Suchman, Edward A. *Evaluative Research*. New York: Russell Sage Foundation, 1967.

McGuire, Thomas O. "Decisions and Curriculum Objectives: A Methodology for Evaluation." *Alberta Journal of Educational Research*, XV (March, 1969), 17-30.

Taylor, Peter A., and Thomas O. McGuire. "A Theoretical Evaluation Model." *Manitoba Journal of Educational Research*, I (1966), 12-17.

These are only a few examples of the kind of modeling that must be done to understand the nature of objectives in a program. Program objectives are usually broader than are objectives that guide instruction. The relationships between the two must be clear.

KINDS OF OBJECTIVES
OPERATING WITHIN
A PROGRAM

Instructional models limit themselves only to objectives dealing with learners behavior. However, a program may have additional kinds of objectives. For

example, programs usually include: *results objectives*, *teacher objectives*, and *administrative or program maintenance objectives*.

Results Objectives. May be of several types. Two of the most common are: (1) learner-centered objectives that indicate what change is to take place in the learners as a result of instruction, and (2) society-centered objectives that specify what will happen to society as a result of program.

Teacher Objectives. These objectives state what the teacher is going to do. They're usually the other side of the coin from the learner objectives. They include such things as: helping the student distinguish between two methods of work or motivating the least responsive members of the group to take an active part.

Administrative or Program Maintenance Objectives. These objectives are institutional in nature, are concerned with program maintenance or instrumental objectives that focus on the mechanism that provides the program.

This category includes objectives dealing with improved procedures, increased enrollments, improved contextual features and personnel.

All three types of the objectives are legitimate and should be evaluated for their attainment.

Learner-centered objectives may be further examined in terms of their origin. Sources may include:

1. Objectives set by the learner himself.
2. Objectives set by the representatives of the expected learners.
3. Objectives recommended by those who know the expected learners very well.
4. Objectives recommended by a larger society or a particular pressure group who want to see another group make certain changes.
5. Objectives set by a subject-matter expert in terms of what the learner should know about the subject.
6. Objectives set by the "think group" of an agency to carry out the expected mission of the agency.
7. Objectives set by teachers from thin air or drawing from any of the other sources.

Just because an objective is phrased in terms of the learner doesn't mean that it's an objective that's originating with any of the learners who actually take part. Supposedly their act of participation means that they accept the objectives that have been announced. However, that's not always the case.

There may be several different and conflicting sets of learner objectives functioning in any learning situation. Some of them will be stated. Others won't. For example, the agency head may think that one set of learner's objectives is guiding the program. The teacher may have developed a set of learner's objectives based on her experience with the discipline, and the learners themselves may have quite heterogeneous sets of objectives neither agreeing with what the teacher has set or with each other.

SOURCE: Steele, Sara M., and Robert E. Brack. "Evaluating the Attainment of Objectives: Process, Problems and Prospects." [Manuscript in process.]

ADDITIONAL

REFERENCES: Uhl, Norman P. *Identifying Institutional Goals*. Durham, North Carolina: National Laboratory for Higher Education, 1971.

Peterson, Richard E. *The Crisis of Purpose: Definition and Uses of Institutional Goals*. Princeton, New Jersey: Educational Testing Service, 1971.

CHARACTERISTICS OF OBJECTIVES

Another way of getting a feel for the phenomena that's called objectives is by identifying the characteristics of objectives and trying to see how those characteristics affect their use and the evaluation of their attainment.

Houle defines objectives as the actual intended result of learning and not just a mere formal expression of that intention. He suggests that objectives have the following attributes that affect programming.

1. Objectives are essentially rational.
2. An objective is practical. The ultimate test of an objective isn't validity but achievability.
3. Objectives lie at the end of actions designed to lead to them.
4. Objectives are usually pluralistic and require the use of judgment to provide a proper balance in their accomplishment.

5. Objectives are hierarchial.
6. Objectives are discriminative.
7. Objectives change during the learning process.

Warner refers to the characteristics of intangibility, change, number, continuousness, and remoteness as being barriers to using objectives effectively in managing programs.

SOURCES: Houle, Cyril O. *The Design of Education*.
San Francisco: Jossey Bass, Inc., 1972.

Warner, W. Keith. "Problems in Measuring the
Goal Attainment of Voluntary Organizations."
Journal of Adult Education, XIX (No. 1, 1967).

GROUP 6

RESULTS--EVALUATION OF OUTCOMES AND EFFECTS

OVERVIEW:

How do you see beyond your objectives and look at what else is happening as a result of your programs? A growing number of approaches are dealing with outcomes and effects that include but go further than just looking at results related to the specific objectives of the program. These types of evaluation strategies are concerned with anticipated and unanticipated results, direct and indirect results, and benefits and harmful consequences. They vary considerably in their approach and coverage. Some of the material included in this section "model" the nature of results, some deal with procedures for examining results.

EXAMPLES INCLUDED:

- *Goal free evaluation* compares the results achieved with the original need as a way of establishing program effectiveness.
- In the *zones of results* approach, three sources of information about results are identified and data-gathering devices that can be used most effectively are discussed.
- The *multiple dimensions of program effectiveness* approach identifies various ways effectiveness can be defined.
- The *multiple change approach* classifies kinds of results expected in community development programs and indicates sources of information.
- *Effectiveness of methods* provides an example of how results data were used to examine the "payoff" of various kinds of programming methods.
- *Impact evaluation* looks at the long-range achievements of a program.
- *Social indicators* are used as a means of measuring community progress.
- *Institutional evaluation* provides a framework for evaluating multi-course curricula.
- *Public policy* approach provides a framework for looking at the relationship of programming to public policy.
- *Research models* review a traditional approach to results evaluation.

- *Efficiency examinations* discuss the role of cost-benefit analysis of results.
- *Accountability* deals with the concept of stewardship and identifies related activities.

GENERAL ADVANTAGES:

- Examination that includes a full spectrum of results gives us a better understanding of the actual effects of programs.
- Such approaches deal with the complexity and dynamism of human nature.
- Examining results in a systematic way helps us become more realistic in our expectations of programs.

GENERAL CAUTIONS:

- Because of the extensive range in results, when resources are limited, some means such as original problem or need, mission of the agency, or prediction of the most crucial positive and negative results, must be used to narrow the scope of results data that will be gathered.
- Standards are needed for interpreting what amount of results it's reasonable to expect from various amounts of program inputs with participants of selected characteristics (examples; task oriented compared with socially oriented; less than high school education compared with college graduates).
- Clear patterns of organizing and managing results data must be set up well in advance of collection. Otherwise, the data may become fragmented and not build a clear and coherent picture.

6.1 GOAL FREE EVALUATION

ORIGIN	Evaluation of educational innovations.
ESSENCE	Results of programs are judged against the originating need. Cost of producing those results is considered in relation to costs of alternatives.
USE	Particularly helpful in situations where objectives are not clear or are unrealistic. Helpful too as a means of examining objectives to see if they limited the potential of the program, that is, is it able to produce more than was originally expected from it?
STARTING POINT	Examine a profile of the needs that the program is addressed to.
DESCRIPTION	One of the more often talked about and controversial models emerging in education is <i>goal free evaluation</i>, GFE. Such evaluation starts without referring to the objectives of the program. If they are used, they're used after the data are in to see which of the results that emerged actually were intended. The base for determining what data to collect is a profile of identified needs. Actual effects are then compared with these needs rather than with the objectives. The model emerged through experience in evaluating products of research and development centers where it was discovered that sometimes the side effects were of greater value than the intended effects. The work of evaluation is viewed primarily as an activity of condensation that includes two major stages--<i>compression</i> and <i>credentialing</i>. A mass of data and observations about a program are condensed through various means (compression) until judgments as to the value and worth of the program can be made (credentialing). The steps in the process have been identified as follows: 1. Characterizing--explore questions like: What is it? What's the situation? What's actually occurring? 2. Clarification--specify type of conclusions wanted: effects? Comparative effects? 3. Causation--dealing with effects. 4. Comprehensive check of consequences--pick up side effects. 5. Criteria--qualities that are meritorious.

6. Costs--was it worth money that went into it? Is it more or less expensive than other ways of getting the same results?
7. Critical comparisons and competitors--do cost analysis on them also.
8. Credentialing--refer back to need assessment, attach merit, inferential step, putting stamp of approval on it.
9. Conclusions--as identified in #2.

Evaluation of effects may be concerned with any or all of the following:

Immediate Performance--Changes In

1. Knowledge.
2. Comprehension or understanding.
3. Motivation (attitudes, values, effect).
4. Nonmental abilities (perceptual, psychomotor, motor, social skills).
5. Long-range changes.

Secondary Effects On

1. The teacher.
2. The teacher's colleagues.
3. Other students.
4. Administrators.
5. Parents.
6. The school.
7. The taxpayer.

If, when the data are examined, results that coincide with the stated objectives are stronger and more conclusive than other results, this is particularly good evidence that the program was well designed to attain certain objectives. However, other results aren't lost in the process.

Evaluation is seen as an activity by which performance data are gathered and combined with a weighted set of goal scales to yield either comparative or numerical ratings, and as the justification of the instruments, goals, and ratings. Cost analysis is essential to explain the basic question of whether it was worth the resources that went into it.

The goal of such evaluation is the answering of questions about value or worth. Usually, this goal is accomplished more effectively if comparisons are made between alternative programs rather than by comparing participants with those who haven't participated in any program of this type.

SOURCES

Scriven, Michael S. "The Methodology of Evaluation." In *Perspectives of Curriculum Evaluation*, Ralph W. Tyler, Robert Gagné, and Michael Scriven, eds. AERA Monograph Series on Curriculum Evaluation, No. 1. Chicago: Rand McNally & Co., 1967.

Scriven, Michael. "Prose and Cons About Goal-Free Evaluation." *Evaluation Comment* (December, 1972), III.

Notes from AERA Concepts of Evaluation Workshop, Portland, Oregon, October 19-20, 1972.

6.2 ZONES OF RESULTS

ORIGIN	Adult education.
ESSENCE	Results need to be examined within the learning setting, in real life during the course of the learning, and in real life at an interval after the completion of the learning activity.
USE	Helpful in identifying sources of information.
STARTING POINT	List the kind of results expected at various times.
DESCRIPTION	<i>"Formative evaluation"</i> discovers whether the instructional methods are producing the desired result on course, providing information on what works with whom and what additional training is required. When a prototype is being developed, it reveals what revisions in the course must be done to produce the desired result (for example, lesson changes, sequencing changes, added lessons, substituted lessons, needed training techniques, etc.) and other modifications while the course is going on. <i>"Summative evaluation"</i> provides information on the total impact of the course. This model for examining the results of a program identifies three main zones that are relevant in programs for adults. They are: Zone 1--Study of Results As Evidenced Within the Class or Program Activities. Simulation, role playing, and gaming are used to trigger evidence of results within the actual program. Zone 2--Application in Real Life. Information is secured about how adults are using information in real life as the program of course progresses. Zone 3--Follow-Up Study. Interview and evaluation procedures are used at various periods after the course. The extent of success of transfer to real life while in the course and after the course should be compared. This approach is particularly interested in the development of programs that will be distributed and used in many locations and as result focus on results at various stages in the development process. The following questions are suggested as a guide to further development.

- a. Which of the behavioral changes specified for the students actually occur as a result of the course?
- b. What behavioral changes actually occur as a result of the course, but aren't specified in the objectives?
- c. In what respects does the course fail to provide the opportunity necessary for the students to achieve the specified behavioral changes?
- d. What changes specified for the course are a function of time (no training)?
- e. Does academic upgrading only result in the same behavioral changes as those specified in the course combined with a sequential course?
- f. Assuming that students have expectations for a "traditional learning setting," will the provision of such a setting result in more efficient development of specified behavioral change?
- g. How do the techniques of skill training produce specified behavioral changes more efficiently than those techniques in the present course?
- h. In what ways does the precision required for the implementation of the techniques of skill training refine the definition of the behavioral objectives in the course?
- i. In what ways do the written lesson materials fail to give adequate guidance to the coach for effective lesson implementation (timings, sequence of lessons, clarity of direction, completeness of direction)?
- j. What skills do the coaches lack to objectively achieve behavioral change in students?

The model describes a variety of kinds of data and means of securing data in relation to one or more of the zones. Each description of a data collection approach has four parts: purpose, criteria, decisions, and procedure, as illustrated in the following example that discusses direct *observation* as a method of securing data.

Purpose. To identify behaviors that occur in response to specific lessons, teachers, and other student behavior.

Criteria. The behaviors must be observable: word, movement, facial expression, body attitude, associations with other persons in the learning group. The observer

doesn't rate the quality of a behavior; he notes only its presence and in the report offers an interpretation of the behavior.

Decisions.

1. The teacher may use these records to modify his teacher behaviors to bring about desired changes in his group. He uses them immediately.
2. The lesson developer uses the records to substitute new directions in the written lesson materials or to add new directions if he can establish the relevance between the noted behavior and the lesson.
3. The observations provide the empirical basis for the description of the theoretical course model. These observations will determine the retention and modification of the model.

Procedure. All training groups are observed through all lessons by trained observers. A behavior analysis that reflects behaviors expected in the lesson is used as a guide by the observer. He records a summary of behaviors exhibited by each group member at the end of each session.

SOURCE

Lamrock, A. L., A. D. Smith, and P. W. Warren. "Evaluation: Its Scope and Systems for Evaluation Development." Paper prepared for meeting of research directors of NewStart Corporations, Ottawa, Saskatchewan, March, 1971.

6.3 MULTIPLE CHANGE APPROACH

ORIGIN	International development programs.
ESSENCE	Four kinds of results, six dimensions, and four major sources of information are outlined.
USE	Particularly useful in programs with goals that go beyond the particular participants.
STARTING POINT	Determine the type(s) of results that can be expected from the program.

DESCRIPTION

This model was designed for examining international development programs.

It suggests that several kinds of changes are needed for development to occur--*psychological changes* in motivation and attitudes; *educational changes* in knowledge, concepts, and skills; *technological changes*, new methods of production, organization, administration, and distribution; and *sociological changes*, relations among persons or groups.

The model distinguishes between first approach or the immediate objective of the program, and more fundamental changes brought about as a result of the first approach. For example, a project that focuses on use of fertilizer might be evaluated only on whether the farmer applied more fertilizer this year. Progressively more fundamental, however, might be the following types of changes that might come from the program. Does he buy fertilizer at commercial prices in subsequent years? Is he more willing to try other new practices? Is he more willing to support or cooperate in the development of new agricultural techniques?

Kinds of results sought by programs can be classified as follows:

Impersonal results--legislation adopted, a remedy for a disease developed.

Administrative results--policies or procedures changed, institutions strengthened.

Results involving persons:

1. Changes in information, attitudes, and skills.

2. Actions and habits of individuals or groups.
3. Changes in social structure or relationships.

Changes in economic and social conditions resulting from actions taken by individuals or groups.

Changes in capital resources.

The standards by which development projects should be judged successes or failures may be considered in terms of six major dimensions:

1. The kind of results sought (including their permanence and spread).
2. The side effects produced.
3. The area of activity and the area of influence where the results appear.
4. The time period within which results of specified size are obtained--the pace of change.
5. The cost of bringing about results of specified size--the efficiency of the project.
6. The amounts of desired results brought about by factors other than the development project.

One of the challenges is selecting which changes will be studied in light of the significance of those changes and the availability and reliability of data related to the changes.

Although the intention and specific objectives of the project are a starting point, an attempt should also be made to predict and examine potential side effects.

Four major types of data are suggested: (1) *already available data*, (2) *additional records* that organizations or individuals might agree to keep, (3) data obtained by *systematic observation* of behavior, and (4) data obtained by *systematic questioning*.

The obtaining of measurements of project results and costs isn't in itself evaluation. They have to be interpreted. *Interpretation is the component that distinguishes evaluation from measurement.*

SOURCE

Hayes, Samuel P., Jr. *Evaluating Development Projects*. Place de Fontenoy, Paris: United National Education, Scientific & Cultural Organization, 1959.

6.4 MULTIPLE DIMENSIONS OF PROGRAM EFFECTIVENESS

ORIGIN	University Extension.
ESSENCE	At least six dimensions of program effectiveness may need to be considered-- <i>general results, attainment of objectives, intent of the program, impact and significance, contribution to mission, and use of resources.</i>
USE	Particularly useful in selecting results data that may be influential on outside audiences.
STARTING POINT	Review of the mission and intent of the program.
DESCRIPTION	All programs need to be examined in terms of their effectiveness. <i>Effectiveness deals with their degree of success in producing results.</i> Effectiveness may have several components. One model proposes that there are at least six dimensions of effectiveness that should be applied to programs: 1. <i>General results.</i> 2. <i>Attainment of objectives.</i> 3. <i>Intent of the program.</i> 4. <i>Impact of significance.</i> 5. <i>Contribution to mission.</i> 6. <i>Use of resources.</i> The two dimensions that aren't covered by the other models in this section are contribution to mission and the intent of the program. Contribution to Mission. Most agencies and units within agencies either are commissioned by legislation or other enabling acts to a certain mission, or develop an overall mission within which they operate. The results of program should be analyzed in terms of the extent to which the mission is actually being accomplished.

Intent of Program. Most adult education programs can be categorized into certain types. Some have the traditional intent of diffusing knowledge and helping people to master content. Others are devoted to problem coping. Still others are concerned with developing generalizable skills in people such as abilities in analysis, creativity, or ability to communicate effectively. Still others are intent on changing conditions. Ultimately, a program's results should be examined in terms of the basic intent of the program. If the intent of the program is problem solving or coping, the basic examination should be of what resulted in terms of the problem. Was it solved, eased, or controlled? If so, the program was effective in realizing the intent of the program. If a lot was learned but nothing done in terms of the problem, the program may have been successful on other dimensions of effectiveness, but failed in terms of its intent. Although things in addition to the intent may have considerable value, part of the definition of success must lie with the overall purpose for which it was undertaken.

SOURCE

Steele, Sara M. *Six Dimensions of Program Effectiveness*. Madison, Wisconsin: University of Wisconsin-Extension, Division of Program and Staff Development, 1972.

6.5 EFFECTIVENESS OF METHODS

ORIGIN	Cooperative Extension.
ESSENCE	Data as to practices used and kinds of contacts with Extension were correlated to determine the effectiveness of various Extension methods.
USE	Useful when there's quantitative results data and clearly distinguishable methods.
STARTING POINT	Examine the nature of practices and results to see if such a comparison can be made.
DESCRIPTION	Cooperative Extension took a nationwide approach to the examination of the comparative effectiveness of various teaching methods during the 1940s. Although the procedure hasn't been used recently except in one or two states, the basic pattern is presented because of its adaptability. First, data were gathered on Extension's influence on the use of farm and home practices. Data were gathered from respondents from more than 15,000 farms and homes in 32 sample areas in 27 states. Only the practices that respondents could associate with Extension teaching were included. Data were then analyzed to determine the number of practices that a particular method, <i>demonstrations, general meetings, office calls, bulletins, etc.</i>, had influenced. Further analysis examined the relationships of methods to content of the practice.
SOURCE	Wilson, Meredith C., and Gladys Gallup. <i>Extension Teaching Methods</i> . Extension Service Circular 495. Washington, D.C.: Federal Extension Service-USDA, August, 1955.

6.6 IMPACT EVALUATION

ORIGIN	Manpower training programs.
ESSENCE	Impact evaluation looks at the ultimate effects of programs.
USE	Particularly useful in examining the extent to which an agency is accomplishing its mission.
STARTING POINT	Examine the mission of the program and identify the major results that program is expected to produce.
DESCRIPTION	Developed as a means of evaluating manpower projects, this model is projected by its authors as one of three approaches to evaluation. The other two are <i>project monitoring</i> that examines the efficiency of a specific project and <i>evaluation of the immediate success or failure</i> of a particular program. In comparison to them, <i>impact evaluation</i> examines the long-run goals of the program and views success and failure in these terms. The goal of such evaluation is to provide policy makers with the basic data necessary for them to make decisions wisely. Impact evaluation should provide five essential sets of information: 1. All the data necessary to determine whether a particular type of program should be continued. 2. Indication of which alternative programs provide the greatest gains for a given cost. 3. Information on the component of each program and the mixes of components that are most effective for a given expenditure so that minimum operating efficiency can be achieved. 4. Information for persons with different characteristics so decision makers can determine which individuals are best served by each program. 5. Information about new methods of attack on this type of problem. The types of people who will benefit should be considered when criteria for success are being established. In most projects, two types are always present--the participants and society as a whole. In other instances, other types like employers, government, etc., are beneficiaries. Criteria

of success should be developed from the standpoint of all that stand to gain or lose through the project.

In each instance, the criteria for success need to be developed into specific goals with clear operational criteria.

Another description of impact evaluation describes it as a term used to indicate the broad effects of a program. It focuses on the mark that a program leaves on the people, community, problem situation, and/or conditions within which the participants operate. It isn't directly concerned with specific results of the program, but it's concerned with how these changes affect other people or change conditions which affect people.

Although most programs have some degree of impact, impact evaluation is concerned with examining whether the program has accomplished something of both magnitude and importance.

This type of evaluation is best done on:

1. Large-scale projects.
2. Projects of considerable force, concentrated intensive input.
3. Longitudinal and long-range programs.
4. Combinations of efforts.
 - a. Related program activities over several years.
 - b. Efforts of a team of people.

SOURCES

Borus, Michael E., and William R. Tash. *Measuring the Impact of Manpower Programs: A Primer*. Ann Arbor, Michigan: The University of Michigan--Wayne State University, Institute of Labor and Industrial Relations, 1970.

"Impact Evaluation." Material prepared as a follow-up of a University of Wisconsin-Extension Workshop on Impact Evaluation, Madison, Wisconsin, March, 1972.

6.7 PUBLIC POLICY

ORIGIN	General education.
ESSENCE	Program outcomes should be considered and classified as to whether they're just programmatic or whether they have public policy consequences.
USE	Particularly relevant to programs affecting disadvantaged adults.
STARTING POINT	Examine the effects of education on public policies.
DESCRIPTION	In addition to dealing with programmatic questions, evaluation should contribute to better understanding and dealing with public policy questions. Programs include outcomes relating to public policy when they:

1. Directly or indirectly alter the power relationship between the citizen and the state.
2. Affect immediately or in the long run the status a person has and the power he can exercise within the social system.
3. Increase or decrease social tensions.
4. Effect a change in the self-concept or sense of self-worth of the individual.

The model identifies ways of categorizing outcomes in terms of the following table:

Realms of Outcomes	Intended Outcomes	Unintended Outcomes	
		Anticipated	Unanticipated
Public policy			
Programmatic			

Intended public policy outcomes are similar to the broad goals of education. *Programmatic intended outcomes* are the specific teaching objectives. Usually, the outcomes are *intended*. However, often programs also produce *unintentional outcomes*. Some of them, although not set as objectives in advance, may be *anticipated*. Others are completely unexpected.

Public policy questions relate to *moral values*. Evaluation that deals with public policy questions must

be able to provide information related to the moral value questions that are involved. However, models for dealing with moral issues are lacking.

SOURCE

Berlak, Harold. "Values, Goals, Public Policy and Educational Evaluation." *Review of Educational Research*, XL (April, 1970), 261-78.

6.8 INSTITUTIONAL EVALUATION

ORIGIN	Higher education.
ESSENCE	Performance on standards of student achievement and attitudes toward the program are key elements in examining the curriculum offered by a department or unit.
USE	Particularly useful when a sequence or package of courses and activities is involved.
STARTING POINT	Establish the general standards that graduates should achieve.
DESCRIPTION	Developed as a framework for evaluating programs in higher education, <i>institutional evaluation</i> is concerned with questions raised by administrators, suppliers of funds, accrediting committees, and internal and external critics. Questions like: Do the objectives underlying the program adequately reflect the institution's goals? Does the benefit achieved by the program justify the expenditure of resources? Are the institution's objectives satisfactorily met; if not, should the institution demand modification of the program or should it modify the goals? In short, is the program acceptable to the institution? On the other hand, <i>project evaluation</i> takes place within the process of curriculum design and deals with questions posed by the originating professor or design team. Both project and institutional evaluation are concerned with the performance of students. However, project evaluation is concerned with performance in terms of the objectives of the course while institutional evaluation is concerned with performance in terms of general standards for the particular type of student. Institutional evaluation generally focuses on the program as a whole rather than on its constituents. It considers the achievement of any particular program in relation to a network of other programs and goals, a function that can't be done at the project level. Institutional evaluation uses two major kinds of data: <i>information on achievement of standards of student achievement</i> and <i>information about attitudes toward the program</i>. A constructive system of university standards is based on the following premises:

1. Standards are subject to continual revision. They must be sensitive to changes within fields of knowledge and in educational philosophy.
2. Standards result from a process of analysis that includes input from a wide range of opinions and perspectives within an institution.
3. Standards are developed in cognizance of, though not necessarily in agreement with, the objectives reflected in the design of relevant programs.

Standards may be best set by committees that include members of the department originating the program, members of the departments where proficiency is needed for their own course work, and other knowledgeable faculty members serving at large. The committee would:

1. Study objectives and procedures of a course being developed.
2. Examine criteria that define what the particular type of student ought to be able to do in the subject area (opinions of relevant faculty members, opinions of professionals outside the organization, specifications developed by professional associations, requirements of upper level courses, requirements of employment).
3. Develop a statement of institutional standards in two forms:
 - a. A conceptual statement identifying concepts and skills that are important parts of achievement.
 - b. An operational statement including standards of achievement and the way in which they will be examined.

The standards examination wouldn't be a course examination used to evaluate the performance of individual students, but would be used periodically to check the performance of the courses of the department.

Attitudes Toward Program. Reactions of students and faculty serve to indicate the degree of a program's acceptance. Opinion polling methodology may be used to secure opinions from students who have participated in the program, students scheduled to participate, and faculty of relevant departments. Opinions about the objectives of the program, effectiveness in meeting the goals, and opinions about continuation of the program would be gotten.

Integration of Information. Curriculum reports should be done with the same care as research reports.

An evaluation study doesn't itself constitute the evaluation. Data must be interpreted and acted on by appropriate representatives of the organization. Evaluation isn't a single study or series of studies; it's an ongoing activity designed to provide answers to questions that arise from a number of sources and perspectives. The number of types of studies conducted at a given time depend on the institution's needs and resources devoted to evaluation. Selection of appropriate studies is an important part of the work of an evaluation program.

The information will be useful only if it does in fact contribute to the institution's appraisal of its own effectiveness. Judgment by responsible organizational members is required. Curriculum evaluation requires the participation of persons from all parts of the institution. An institution-wide Curriculum Review Committee is recommended.

SOURCE

Forehand, Garlie A. "An Evaluation System for Curriculum Innovation." *Teacher's College Record*, LXXII (May, 1971), 576-91.

6.9 SOCIAL INDICATORS

ORIGIN	Social planning.
ESSENCE	Programs dealing with visible entities that are recorded in public data can use those indexes to plan and evaluate programs.
USE	Use is restricted to those types of programs where it's possible to use social-indicator type data.
STARTING POINT	Identify the essential attributes that are within the scope of the program.
DESCRIPTION	Those agencies carrying out broad-based, many-faceted programs in a community might want to consider measuring results in terms of social indexes. The Community Health model, for example, uses measures on indexes in political, environmental, economic, social, and educational fields to analyze community situations as a basis for planning and uses follow-up measures as a means of examining results. Each of the major areas is examined by means of a group of subscales. For example, one of the subunits in the political area is participation in the electoral process. It's examined in terms of such data as the percentage of the adults registered to vote, the percentage that really do vote, participation by ethnic groups. A process for using social indicators includes: <i>Step 1.</i> Identify the essential variables that describe the whole system of reality essential to understanding the meaning of a healthy community. <i>Step 2.</i> Identify the essential factors or subvariables that constitute a reliable description of each variable. <i>Step 3.</i> Determine the most appropriate forms for measuring them--that is, determine which social indicators will be used. <i>Step 4.</i> Develop norms or targets as standards for defining a healthy community. <i>Step 5.</i> Determine how citizens in the community define well being in the context of the variables and subvariables.

- Step 6.* Make a comprehensive examination of the community's own reality.
- Step 7.* Analyze existing public and private programs.
- Step 8.* Determine what should be done.
- Step 9.* Determine the costs of alternative programs.
- Step 10.* Assign program responsibility.
- Step 11.* Design mechanisms for evaluating the program so that there can be regular assessment of the relationship between the normative standards that the community has set for itself and the existing reality.

SOURCE

Paulson, Belden. *A Model for Community Analysis: Steps in Planning the Total Health of a Community*. Milwaukee, Wisconsin: University of Wisconsin-Extension, Center for Leadership Development, April, 1972.

ADDITIONAL
REFERENCES

Munson, Byron E. *Changing Community Dimensions*. Columbus, Ohio: Ohio State University, 1968.

Etzioni, Amitai, and Edward W. Lehman. "Some Dangers in 'Valid' Social Measurement." *The Annals of the American Academy*, II (September, 1967), 1-15.

Toward a Social Report. Washington, D.C.: U.S. Department of Health, Education and Welfare, January, 1969.

6.10 RESEARCH MODELS

ORIGIN	Social science.
ESSENCE	The presence or absence of results can be detected by using standard research design.
USE	Of limited value in situations where program input is too low to produce major measurable results or external variables are complex and not subject to research controls.
STARTING POINT	Establish the dimensions of the study.
DESCRIPTION	Some evaluation models duplicate research models with few if any modifications. Three research approaches to evaluation are given as examples: Longest suggests the following process: 1. Define the general conditions and problems that require the evaluation to be made. 2. Specify theories, assumptions, and values that are accepted as true and will guide the formulation of the evaluation research. 3. Specify exactly what it is that will be evaluated. 4. Formulate hypotheses. 5. Identify variables. 6. Specify units of analysis. 7. Collect data. 8. Analyze data. 9. Plan for reporting. 10. Plan for executing. Weiss identifies the following core elements: 1. Formulating the program goals that evaluation will use as criteria. 2. Choosing among multiple goals.

3. Investigating unanticipated consequences.
4. Measuring outcomes.
5. Specifying what the program is.
6. Measuring program inputs and intervening processes.
7. Collecting the necessary data.

She deals with selecting crucial factors from among dependent, independent, and intervening variables. The dependent variables are the indicators of the program outcomes. The inputs are the independent variables. The intervening variables are the factors that mediate between the two. Adequate indicators usually involve multiple measurement.

Outcome variables may include such things as:

Effects on persons served.

Effects on agencies.

Effects on larger systems.

Effects on the public.

Input variables include:

Program variables: purpose, principles, methods, staffing, persons served, length of service, location, size of program, auspices, management.

Participant variables: age, sex, socioeconomic status, race, length of residence in community, attitudes toward the program, motivations for participation, aspirations, expectations from the program, attitudes of other family members, degree of support from others for the intent of the program.

Intervening variables include:

Program-operation variables.

Bridging variables: attainment of intermediate milestones.

Developing a picture of the intended processes of a program may be helpful in examining which variables to use or control.

Although it uses program goals as a central focus, it also recognizes the importance of examining unanticipated consequences.

The foregoing examples apply a deductive research approach. A third, less well-formalized approach draws on inductive approaches to research. Tentatively labeled "search" evaluation by Cain and Hollister, it sets out to determine what happens during and as a result of a program. However, the authors are more prone to recommend use of intentional experiments.

SOURCES

Longest, James W. "Designing Evaluative Research." Paper presented at the annual meeting of the Rural Sociological Society, San Francisco, California, 1969.

Weiss, Carol H. *Evaluation Research*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1972.

Cain, Glen G., and Robinson G. Hollister. *The Methodology of Evaluating Social Action Programs*. Madison, Wisconsin: University of Wisconsin, Institute for Research on Poverty, n.d., pp. 9 and 42-69.

6.11 EFFICIENCY EXAMINATIONS

ORIGIN	Industry and the military.
ESSENCE	Evaluation of the amount of benefits in comparison with the costs of the benefits and with alternative investments of the same resources.
USE	Particularly useful when time, money, or energy resources of the participants or the programming agency are limited.
STARTING POINT	Develop units that describe the costs and the benefits of a program.
DESCRIPTION	Efficiency includes results but goes beyond them to examine the cost of those results. Examination of efficiency usually involves some type of cost-benefit analysis. Costs may be of two kinds: <i>direct costs</i> and <i>opportunity costs</i>. Opportunity costs consist of other benefits forgone. Benefits also are of at least two types. The most commonly mentioned types are <i>private</i>, those accruing to the individuals involved, and <i>social</i>, those that accrue to society and not to one individual. There are usually five elements present in a cost-benefit analysis: (1) the objective (or objectives), (2) alternatives for reaching those objectives, (3) costs, (4) a model (or models), and (5) a criterion. The criterion is the rule or standard by which the alternatives are ranked in order of desirability and the most promising ones chosen. The process involved includes the following steps: 1. Identifying the pertinent measures of effectiveness, that is, benefits. 2. Describing the alternatives. 3. Expressing both mission performance and cost as functions of the characteristics of each alternative. 4. Estimating appropriate values for the equation parameters. 5. Computing, analyzing, and presenting results. In some instances, cost-benefit is concerned only with the proportion of costs to benefits for the given program approach. More often, however, the approach is more useful if it deals

with alternative usage. It explores such questions as: Is there any better way to attain the same results? Can the same end result be achieved at less cost? Which of two approaches will produce most at least cost? Would there have been a greater value to the participants, to the agency, or to society if the same resources had been invested in a different program?

Adult education programs need to consider the adult's resources and costs to participants in terms of time and energy, and the resources of society that are being expended if the particular program is tax supported, as well as the usual concept of agency resource inputs.

Some of the kinds of program outputs that may need to be considered are:

1. *Economic improvement*--gain in economic status of the participants, community, group, or region.
2. *Action*--outputs of direct action.
3. *Specific knowledge, attitude, or skill.*
4. *Development*--personal or group.
5. *Enhanced self-concept.*
6. *Satisfaction and enjoyment.*
7. *Prestige.*

Often outputs need to be transferred into some quantitative unit other than dollars. This is sometimes done by establishing a symbolic value unit and measuring the quantities of such units generated by a particular program.

SOURCE

Steele, Sara M. *Cost-Benefit Analysis and the Adult Educator*. Syracuse, New York: ERIC Clearinghouse on Adult Education and Adult Education Association of the U.S.A., 1971.

6.12 ACCOUNTABILITY

ORIGIN	General education.
ESSENCE	Programmers funded through tax monies are stewards of those monies and responsible for producing adequate results.
USE	Accountability receives greater stress the farther removed the program is from its source of funds and/or the larger the groups involved.
STARTING POINT	Analyze who your program is accountable to.
DESCRIPTION	Accountability is a process occurring between those entrusted with the accomplishments of specific tasks (<i>stewards</i>) and those having the power of review (<i>reviewers</i>). The heart of the process is for the party "standing to account"--the steward--to explain as rationally as possible the results of efforts to achieve the specific task objectives of his stewardship. Reviewers will be concerned with the matching of performance and attainment levels against their expectations as expressed in task specifications and making a determination of their level of confidence in the steward and his efforts. The steward standing-to-account will be concerned with his ability to accomplish the specified tasks and his ability to explain attainment levels in a manner that maintains or builds the reviewer's confidence in his stewardship. Various approaches to accountability are being advocated. One involves specific performance contracting on the part of teachers. Another increases the autonomy of local schools and builds increased local participation. However, instead of, or in addition to, such specific activities, improved accountability is seen as having 12 critical factors. Community Involvement. The use of members of concerned community groups in appropriate phases of program activity to facilitate: program access to community resources; community understanding of the program's objectives, procedures, and accomplishments; and the discharge of program responsibilities to relevant community, client, service, and support groups. Technical Assistance. The means for providing adequate resources in program planning, implementation, operation, and evaluation by drawing on community, business, industrial,

labor, educational, scientific, artistic, social/welfare, and governmental agencies for expertise and services necessary to effective operations.

Needs Assessment. The identification of target-group and situational factors essential to the planning of a relevant program of action.

Change Strategies. The development of effective strategies for systematic change in the educational enterprise and the incorporation of the strategies of program operations.

Management Systems. The adaptation of the systems approach, through such techniques as Management by Objectives, PPBS, PERT, CPM to educational program management at the local, state, and federal levels.

Performance Objectives. The specification of program objectives in a comprehensive, precise manner that indicates measures and means for assessing the degree of attainment of predetermined standards.

Performance Budgeting. The allocation of fiscal resources in accordance with program objectives to be realized, rather than by objectives or functions to be supported.

Performance Contracting. The arrangement for technical assistance in program operations through internal or external contracts where compensation is based on the accomplishment of specified performance objectives.

Staff Development. The determination of the nature and extent of staff development needed for the successful implementation of the accountability concept at the local, state, and federal level, and the design and conduct of indicated development activities.

Comprehensive Evaluation. The establishment of systems of performance control based on the continual assessment of the program's operational and management processes and resultant products.

Cost Effectiveness. The analysis of unit results obtained in relation to unit resources consumed under alternative approaches to program operation, as a determinant in continued program planning.

Program Auditing. A performance control system based on external reviews conducted by qualified outside technical assistance, designed to verify the results of the evaluation of an educational program and to assess the appropriateness of program operation and management.

The search for new ways of implementing the accountability concept in education can be expected to continue. The current range of ideas, practices, and definitions of educational accountability is broad and differing in rigorousness. But the pressure of the times and the advances of technology should be enough to ensure the continuation of the search.

SOURCE

Browder, Leslie H., Jr. *Emerging Patterns of Administrative Accountability*. Berkeley, California: McCutchan Publishing Corporation, 1971, pp. 1-21.

**ADDITIONAL
REFERENCE**

"Accountability in Education." *Educational Technology* XI (January, 1971).

Appendix A

CONDENSED CONTEMPORARY PROGRAM EVALUATION APPROACHES

Appendix A

CONDENSED CONTEMPORARY PROGRAM EVALUATION APPROACHES

Below are condensations of contemporary program evaluation approaches.

The upper-left corner includes the title and author of the approach. The upper-right corner lists the group number and page number where the approach is located in the text. The approaches are alphabetized by title.

Accountability

6.12
p. 217

Accountability is a process occurring between those entrusted with the accomplishments of specific tasks, stewards, and those having the power of review, reviewers. The heart of the process is for the party "standing to account"--the steward--to explain as rationally as possible the results of efforts to achieve the specific task objectives of his stewardship.

Actual Component Approach

Knox, Mezirow, and Darkenwald, Jr.

2.2
p. 104

Evaluation systematically examines qualitative differences in program expectations of administrators and supervisors, teachers, students, and among these three groups by analyzing incongruities between intent and reported practices. Evaluation is focused on: *staffing, recruitment, collaboration, in-service education, goals, instruction, and the relationship among these components*. Each is seen as having *inputs, processes, and outcomes*, with the outcome of instruction being the ultimate focus.

Appraisal Model

Harris

4.1
p. 135

Appraisal is an *act of judgment* in which the judging implies both a *criterion*--a standard of some sort--and a *pertinent description* of what's being done. The criterion and the observation must deal with the same thing. Aspects of a program that are subject to appraisal include: *plans and purposes, resources, processes, and effects*.

Countenance of Evaluation
Stake

3.1
p. 120

Evaluation data are either *descriptive--intents* and *observations--or judgmental*. Both descriptive and judgmental data are gathered about *antecedents, transactions, and outcomes*. Examining *contingencies* and *congruencies* among the sets of data is an important part of evaluation. For example, is what was observed congruent with what was intended?

Criteria and Objectives
Matteson

5.3
p. 174

Evaluation includes: purpose of evaluation, criteria selection, preparation for evidence collection, evidence collection, analysis and interpretation, reporting and use. Criteria areas include: *effectiveness, efficiency, suitability, and importance*.

Criteria of Success Approach
Suchman

5.3
p. 171

Attention is given to a *sequence of objectives* from immediate to intermediate to long range. An understanding of the *assumptions* that underlie the objectives and a checking of the validity of those assumptions is essential.

Five general categories of criteria according to which the success or failure of a program in attaining its immediate-, intermediate-, or long-range objectives may be evaluated are: *effort, performance, adequacy of performance, efficiency, and process*.

Cube Models
Hammond; Armstrong *et al.*

5.3
p. 180

Cube models show the relationship of program to the attainment of objectives. The three sides of the cube are: *behavioral dimension, instructional dimension, and institutional dimension*. Sub-items within the dimensions represent types of variables that may need to be considered in setting objectives and in examining the effectiveness of programs in meeting objectives.

Data Management
Phi Delta Kappa

4.2
p. 138

Most evaluation involves some handling of data. This approach to data management identifies four major elements: (1) delineating the information system needed, (2) obtaining the information, (3) providing the information, and (4) evaluating the evaluation. Soundness of the data rely in part on the soundness of the assumptions that are made in the data-handling processes. Six criteria should be kept in mind if evaluative data are to be informative to the receiver: *relevance, importance, scope, credibility, timeliness, and pervasiveness.*

Decision Centered Evaluation (CIPP)
Stufflebeam; Phi Delta Kappa

1.1
p. 78

Evaluation provides information for judging decision alternatives. It can be useful at all stages of decision making--*awareness, design, choice, and action.* Four types of evaluation--*context, input, process, and product*--correspond to four major kinds of programming decisions--*planning, structuring, implementing, and recycling.*

Differential Evaluation
Tripodi, Fellin, & Epstein

1.2
p. 81

Evaluation is a management technique through which systematic feedback is used in improving programs. Programming is examined in terms of the *efforts invested, effectiveness of those efforts, and efficiency* with which the effectiveness is achieved. In addition to the three types of examinations, three program stages are indicated: *program initiation, program contact, and program implementation.* The term *differential* refers to the fact that the effort, effectiveness and efficiency questions asked are different in each of the three stages. Techniques used in doing evaluation are grouped into three categories: *monitoring techniques, social research technique, and cost-analysis technique.*

Discrepancy Evaluation
Provus

1.3
p. 84

Evaluation identifies *discrepancies* between *actual programs* and *standards* for programs so that programs can be improved. Program personnel set standards for activities and results at each programming stage--*design, installation, process, and product.* Actual performance is compared with the standard and discrepancies or areas for improvement identified. Discrepancy information is used either to change the performance or to change the standard.

Effectiveness of Methods
Wilson & Gallup

6.5
 p. 202

Data as to practices used and kinds of contacts with Extension are correlated to determine the effectiveness of various Extension methods.

Efficiency Examinations

6.11
 p. 215

Efficiency includes results, but goes beyond them to examine the cost of those results. Examination of efficiency usually involves some type of *cost-benefit analysis*.

Costs may be of two kinds: *direct costs* and *opportunity costs*. Opportunity costs consist of other benefits foregone. Benefits also are of at least two types. The most commonly mentioned types are *private*, those accruing to the individuals involved, and *social*, those that accrue to society and not to one individual.

Evaluating Decision Making
Hesseling

2.3
 p. 115

This approach suggests that a matrix consisting of: (1) the *problem* requiring the decision, (2) a list of *alternative decisions*, and (3) a generic *list of criteria* used to consider each alternative decision be used as a means of analyzing the appropriateness of the actual decision made. Criteria are classified as: *technological*, *sociopsychological*, and *economic*.

Evaluation As Facilitation of Learning
Bloom, Hastings, & Madaus

5.2
 p. 168

Evaluation is a method of acquiring and processing evidence needed to improve the student's learning and the teaching. It's a tool in education practice for determining whether alternative procedures are equally effective in achieving a set of educational ends. A two-dimensional figure shows relationships among *analysis of the learners*, *instructional decisions*, and *evaluation* and *students*, *instructional processes*, and *objectives*. Three stages of evaluation are indicated: *initial*, *formative*, and *summative*.

Execution-Impact Approach
Freeman & Sherwood

2.1
 p. 102

Included in a discussion of research's role in social policy formation, this approach suggests that the policy maker needs to determine whether: (1) the program was carried out according to prescriptions set forth in the planning and development stages--*assessment of execution*, (2) the program worked--*assessment of efficacy of impact*, and (3) the expenditure of resources was the best considering alternatives--*assessment of efficiency of impact*.

Four-Question Approach
Gottman & Clasen

5.1
p. 156

A recent text built around the Tylerian approach uses four questions as a means of organizing evaluation. They are: *Why? What? How? and How will you know?*

Goal Free Evaluation
Scriven

6.1
p. 192

Results of programs are judged against the *originating need* rather than the stated objectives. Cost of producing those results is considered in relation to *costs of alternatives*. The work of evaluation is viewed as primarily an act of *condensation* that includes two major stages--*compression* and *credentialling*. A mass of data and observations are compressed until judgments as to the value or worth of the program can be made--credentials assigned.

Goal-Referenced Instruction
Popham & Baker

5.2
p. 164

Goal-referenced instruction has four parts: *specification of objectives*, *preassessment*, *instruction*, and *evaluation*. Two kinds of techniques can be used to judge the degree to which the learner has attained the instructional objective: (1) *observation* of behavior and (2) *examination of products* that the learner will produce. Evaluation is an act of assessing merit by judgmentally comparing the observed result of some educational enterprise with a desired standard or criterion of acceptability.

Impact Evaluation
Borus & Tash

6.6
p. 203

This model is projected as one of three approaches to evaluation. The other two are *project monitoring* that examines the efficiency of a specific project and *evaluation of the immediate success or failure* of a particular program. In comparison to them, *impact evaluation* examines the long-run goals of the program and views success and failure in these terms.

Improvement Evaluation
Kreitlow

4.5
p. 148

The judgment of outsiders can play an important catalytic role in helping local programs identify areas where improvements need to be made, stimulate action in areas where improvements must be made, and reinforce confidence in those areas where the program excels. Those commissioning outside appraisers need to clearly define the *specific purposes* of the evaluation and establish *priorities* as to what will be examined.

Information Domains
Nelson

5.3
p. 181

This model suggests three main domains--*ingredients, processes, and product*--from which information should be gathered in terms of how a curriculum attains its objectives.

Institutional Evaluation
Forehand

6.8
p. 207

Institutional evaluation is concerned with questions raised by administrators, suppliers of funds, accrediting committees, and internal and external critics. Questions like: Do the objectives underlying the program adequately reflect the institution's goals? Does the benefit achieved by the program justify the expenditure of resources? Are the institution's objectives satisfactorily met? If not, should the institution demand modification of the program or should it modify the goals? In short, is the program acceptable to the institution?

IPI Formative Evaluation
Lindvall & Cox

1.5
p. 89

Formative evaluation is the *continuous evaluation* of *all elements* of a developing educational program as an aid to the development process. It deals with the basic question: How can every element and operation in the program be examined so it contributes to its improvement? Subquestions deal with the goals to be achieved, the plan for achieving the goals, the operating program, and the achievements of the program.

Locally Directed Evaluation
Byram & Robertson

5.3
p. 176

Emphasis is on involving *local people (staff and citizens)* in making evaluation of the extent the program is reaching its objectives. Essential elements include: commitment by local administration, strategically placed local leadership team, active staff steering committee, functioning citizens' committee, allocation of time to the leadership team, training in research and evaluation procedures, communications and visibility, qualified consultant service.

Macro System Model *Alkin*

2.3
p. 110

Evaluation is the process of first identifying and then quantifying the *relationships between student inputs and educational outputs*, given a constant financial input and controlling for effects of external systems. Components include: *student inputs*; *educational outputs*--changes in students and the impact of the program on home, community, other programs; *financial inputs*, *mediating factors* or the way the financial inputs invested in personnel, organization, instructional design, etc., are used in the program in combination with student inputs; and *external systems*--social, political, legal, economic outside of the school that encompass the program and have impact on it.

Management Information Systems (SEMIS)

2.3
p. 114

The Extension Management Information System is an example of computerized systems for pooling program data. The Extension system consists of records of *planning*, *activity reporting*, and *evaluation*. The three sets of data are used as a basis for administrative and professional decisions.

Means-Ends Hierarchy *Bennett*

3.3
p. 129

The model shows the relationship of seven levels of evaluation data: *inputs*, *activities*, *people involved*, *reaction*, *"KAS" changes*, *change in practices*, and *ultimate results*. The means-ends hierarchy provides a framework for identifying specific evaluation tasks and seeing them in the perspective of other tasks.

Monitoring Evaluation *Bruce*

4.4
p. 146

Monitoring evaluation deals with the *critical management tasks* in keeping programs on track. It's continuous rather than episodic. Major elements of the program included in the model are: *time*, *program action*, and *expected traditional outcomes*. Evaluation cycles consist of four stages; *design*, *observations*, *judgment*, and *adjustment*.

Motivational Model *Lewis*

2.4
p. 118

Motivation helps to explain human behavior. Analysis of motivation may be important in program evaluation. Education is concerned both with motivation within (professional) and motivation without (the client). In each instance, there's *input*, *motivational processes*, and *output*. Motivational processes are grouped into three categories: *organization*, *group integration*, and *personal*. Evaluative criteria are identified for the various processes.

Multiple Change Approach
Hayes, Jr.

6.3
p. 198

Five kinds of results, six dimensions, and four major sources of information to be considered in evaluation are outlined. Results include: impersonal results, administrative results, results involving persons, changes in economic and social conditions, and changes in capital resources. The six dimensions include: kind of results sought, side effects produced, area of influence, time period within which the results are attained, cost of bringing about the results, and effect of factors other than the development project. The four types of data include: already available data, additional records, systematic observation, and systematic questioning.

Multiple Dimensions of Program Effectiveness
Steele

6.4
p. 200

At least six dimensions of program effectiveness may need to be considered in evaluation--*general results, attainment of objectives, intent of the program, impact and significance, contribution to mission, and use of resources.*

National Assessment Program

5.1
p. 155

The national educational assessment activity provides a model for national evaluation. Data were collected on the extent to which American education was attaining its objectives. The objectives used were those negotiated from among those suggested by *school systems, scholars of the discipline, and discerning lay adults.* Emphasis was on the things that *all individuals* of a certain age should be able to do; what *average youngsters* of that age should be able to do, and things that the *most advanced* can do. *Sampling procedures* were used both in selecting the behaviors that would be examined and in selecting the random group of youngsters to be tested.

Natural Process Approach
Steele

4.3
p. 143

The purpose of evaluation is to help the programmer assess the value of programs and programming activity. Evaluation is an everyday process used by most human beings. The challenge is that of helping humans make a process that is subjective (within the mind) become more accurate. Program evaluation is a generic process in that the same basic activity is involved in evaluating the whole program or any part of it. This process has three elements: *criteria, evidence, and judgments.* Accuracy is increased by improving each of the three elements.

NEA Model
 Taba & Sawin

5.3
 p. 179

In an effort to broaden objectives based model, four major components are graphically presented as a square. Objectives are on the top of the square with evidence of *pupil behavior* pertaining to the objectives on the bottom. Evidence of *teaching-learning operations* and evidence of *nonschool factors* affecting learning complete the other two sides of the square.

NewStart Evaluation System
 Lamrock, Smith, & Warren

1.5
 p. 91

Evaluation is essential in the development of packaged courses. The general scope of such evaluation is indicated by: *topic areas where decisions are made, type of data base used in the decision, source of data, and time of the decision*. Formative/summative evaluation integrates different methods into a total research study.

O-A-R Model
 Deniston *et al.*

5.3
 p. 173

Program performance is evaluated by examining the effectiveness and efficiency of the program. Evaluation of both characteristics focus on the relationship of three key components: *objectives, activities, and resources*. Symbolic patterns are developed by using the first letter of each of the key components. Measures can be attained of program effectiveness, AO:PO; program efficiency, AO:AR; and activity efficiency, AO:AA. P stands for planned and A for actual.

Ohio Model
 Starr *et al.*

5.3
 p. 177

This model was designed on a statewide basis. It includes: *program objectives and measurable goal statements; a data set, instrumentation, and procedures to measure the effectiveness* of efforts to achieve specific program goals; the *means by which the evaluation system would be articulated* with larger management systems of state divisions of vocational education. Attention was given to the time frame required for collecting data, the costs in carrying out the procedure, and the personnel required to manage the data.

Organizational Models
 Etzioni; Schulburg & Baker

2.4
 p. 117

Organizational models deal with an organization as a working unit capable of achieving goals. It includes both *functions direct to the program* and *maintenance and support functions essential to the system*.

Organization As a Total System
Young

2.3
p. 109

The total system needs to be understood to facilitate production. Organization is a process with *inputs* and *outputs*; *management* represents the control or feedback mechanism. *Design criteria* are rules used to evaluate the acceptability of designs. Among the most widely used are measurability, feasibility, optimality, reliability, and stability.

Participant Reaction Approaches

1.7
p. 98

A variety of techniques including steering committees, analysis of attendance and re-enrollment records, and end-of-session reaction forms are used to get participant reactions to programs. Information is secured in terms of reactions to one or more of the following: *content, format, facilities, teacher's performance, perception of value, follow-up needed.*

Priority Decisions
Boyle

1.4
p. 86

Evaluation is important in setting program priorities. Criteria for making decisions about priorities are grouped into six broad categories: *society-community, clientele, political, organizational, resources, and personal.* The specific criteria to be used must evolve from the key influences affecting the particular program and programming situation.

Program Contact System
Duft

2.4
p. 112

This approach uses system analysis as a means of initiating a smooth, effective, and efficient flow of information from those providing it to those in need of it for decision purposes. The programming system includes four phases: *program formulation, decision phase, execution, and evaluation of program results.* The decision phase includes four subphases: *program development, program strategy, program dissemination strategy, and program promotion strategy.*

Program Evaluation and Review Technique (PERT)

2.3
p. 107

Projects are broken into specific task components. *Sequence* and *interface* are identified through graphic presentation. *Responsibility, time estimates, and deadlines* are established. *Key steps* and *pivotal activities* are identified. The flow chart provides a basis for analyzing the progress of work as the project progresses.

Program Planning Budgeting System (PPBS)

5.3
p. 178

PPBS has four major phases: (1) *planning*, which is concerned with generating program objectives; (2) *programming*, which is concerned with the preparation of alternative sets of activities and services to achieve program objectives; (3) *budgeting*, which is involved with the formulation of detailed plans, accounting, and reporting; and (4) *evaluating*, which is concerned with progress outputs and effectiveness of programs.

Public Policy *Berlak*

6.7
p. 205

Program outcomes should be considered and classified as to whether they're just *programmatic* or whether they have *public policy consequences*. Public policy questions relate to *moral values*. Evaluation that deals with public policy questions must be able to provide information related to the moral value questions that are involved.

Research Models *Weiss; Longest; Cain & Hollister*

6.10
p. 212

The presence or absence of results can be detected by using a standard research design.

Social Indicators *Paulson*

6.9
p. 210

Agencies carrying out broad-based, many-faceted programs in a community might want to consider measuring results in terms of *social indexes*. The Community Health model, for example, uses measures on indexes in political, environmental, economic, social, and educational fields to analyze community situations as a basis for planning and then to use follow-up measures to examine results. Each of the major areas is examined by means of a group of subscales.

Social Systems Models *Loomis*

2.4
p. 116

Social systems models may have a good deal to contribute to program evaluation. One model includes three key kinds of ingredients: *elements*, *processes*, and *requisites for social actions*. Elements include: standards, sanctions, knowledge, and attitudes which are key ingredients in evaluation. Processes such as communication, institutionalization, boundary maintenance, systemic linkage, and social control affect both a program and its evaluation.

System Approach to Goal Setting
Van Gogh & Hill

2.3
 p. 111

System analysis is a process, a rational mode of approaching complex problems, a tool in decision making. A systems approach to goal setting includes: *systems definition, structure, and design; system design process; defining the boundaries of the system and identifying how those within the system judge its performance; goal integration and methods of obtaining it; goal-setting process.*

System Role Model
Knox

3.2
 p. 124

Seven components are identified from system and role theory and prior research and evaluation as the key targets of analysis in evaluation. They include: *context, inputs, process, activity, outcome, judgment, and application of findings.* The following points should be considered in designing evaluation systems: evidence, benefits, frequency, feedback, commitment, objectivity, and standards.

Trade-Off and Comparative Cost Approach
Glass

1.6
 p. 95

A format for evaluation of materials or educational activities includes: describing the *product* to be evaluated; evaluating the *goals* of the product; clarifying the *point of entry* of the evaluator; determining the kinds of *trade-offs* involved; comparing *costs* with costs of *alternatives*; making an *intrinsic* (secondary) *evaluation*; making an *outcome* (primary) *evaluation*; forming *judgments and recommendations*; stipulating *circumstances* that would modify the conclusions; and *evaluating the evaluator.*

Transactional Evaluation
Rippey

4.6
 p. 150

Transactional evaluation looks at the *effects* of changed programs on the *changers* themselves. It focuses on evaluating *program acceptance* and gives attention to examining: unanticipated as well as expected consequences, the effect on the total organization and not just those directly involved, dislocation due to competition for resources. Both the *supporters* and the *critics* of the program change are involved.

Tylerian Models

5.1
p. 154

Evaluation is concerned with determining whether education is actually producing the *results it sets out to achieve as indicated by statements of objectives*. The evaluation process includes: analyzing objectives to identify and clarify content and behavior, identifying situations which will give the student a chance to express the behavior related to the content, selecting or developing instruments, analyzing the amount of change that has taken place in students, analyzing the strengths and weaknesses of the program.

Types of Evaluation

5.3
p. 181*Wholey et al.*

An analysis of federal program evaluation came up with four categories of evaluation: *program impact evaluation*, *program strategy evaluation*, *project evaluation*, and *project rating*. Monitoring, reporting systems, and cost analysis are defined but excluded from the categorization. Choice of type depends on importance and feasibility.

Weighted Criteria Approach

1.6
p. 96*Crane & Abt*

Cost effectiveness of alternative curriculum materials is calculated by a detailed breakdown and analysis of *components*, *quality*, and *cost*. The following elements are included: *coverage*, *appropriateness*, *motivational effectiveness*, and *cost*. A score is computed using weights assigned to subcriteria. Scores of alternative products are compared.

When-To-Do-It-Yourself Continuum

3.4
p. 131*Alexander*

Data-gathering activities are grouped into five main categories: *habitual but unorganized*, *simple guides*, *reporting by students or teacher*, *post- or pre- and post-testing*, and *experimental design using control or comparative group*. Although program personnel are responsible for carrying out most of the activities, help should be gotten from specialists when the more complicated categories are involved.

Zones of Results

6.2
p. 195*Lamrock, Smith, & Warren*

Results need to be examined *within the learning setting*, in *real life during the course of the learning*, and in *real life at an interval after the completion of the learning activity*.

Appendix B

EVALUATION APPROACHES CATEGORIZED BY FIELD OF ORIGIN

ADULT EDUCATION

Adult Basic Education

Actual Component Approach [Knox, Mezirow, & Darkenwald, Jr.]
Improvement Evaluation [Kreitlow]

Community Development

Multiple Change Approach [Hayes, Jr.]
Social Indicators [Paulson]

Continuing Education

System Role Model [Knox]
Adoption of the Tyler Model by Adult Education
Participant Reaction Approaches

Extension Education

Priority Decisions [Boyle]
Developmental Evaluation--NewStart Evaluation System [Lamrock, Smith, & Warren]
Motivational Model [Lewis]
Program Contact System [Duft]
Management Information Systems (SEMIS)
Means-Ends Hierarchy [Bennett]
Natural Process Approach [Steele]
Monitoring Evaluation [Bruce]
Adoption of the Tyler Model by Adult Education
Multiple Dimensions of Program Effectiveness [Steele]
Effectiveness of Methods [Wilson & Gallup]
When-To-Do-It-Yourself Continuum [Alexander]

Vocational Education

Criteria and Objectives [Matteson]
Locally Directed Evaluation [Byram & Robertson]
Ohio Model [Starr et al.]
Information Domains [Nelson]
Impact Evaluation [Borus & Tash]

OTHER FIELDS OF EDUCATION

Higher Education

Institutional Evaluation [F. E. Hand]
Materials Evaluation--Trade-Off and Comparative Cost Approach [Glass]

OTHER FIELDS OF EDUCATION (cont.)

Elementary and Secondary Education

Decision Centered Evaluation (CIPP) [Stufflebeam; Phi Delta Kappa]
 Discrepancy Evaluation [Provus]
 Developmental Evaluation--IPI Formative Evaluation [Lindvall & Cox]
 Materials Evaluation--Weighted Criteria Approach [Crane & Abt]
 Macro System Model [Alkin]
 System Approach to Goal Setting [Van Gigh & Hill]
 Countenance of Evaluation [Stake]
 Appraisal Model [Harris]
 Data Management [Phi Delta Kappa]
 Transactional Evaluation [Riphey]
 Tylerian Models and Adaptations [Tyler; Gottman & Clasen; Merwin & Womer]
 Goal-Referenced Instruction [Popham & Baker]
 Evaluation As Facilitation of Learning [Bloom, Hastings, & Madaus]
 Objectives: Categorizations and Criteria
 Goal Free Evaluation [Scriven]
 Public Policy [Berlak]
 Accountability

OTHER SOCIAL FIELDS

Health

Criteria of Success Approach [Suchman]
 O-A-R Model [Deniston et al.]

Several Social Fields

Differential Evaluation [Tripodi, Fellin, & Epstein]
 Execution-Impact Approach [Freeman & Sherwood]
 Managerial Systems--Program Evaluation and Review Technique (PERT)
 Organization As a Total System [Young]
 Evaluating Decision Making [Hesseling]
 Socio-Organizational Systems
 Social Systems Models [Loomis]
 Organizational Models [Etizioni; Schulberg & Baker]
 Types of Evaluation [Wholey et al.]
 Research Models [Weiss; Longest; Cain & Robertson]
 Efficiency Examinations

Appendix C

A TRY AT DEVELOPING A TAXONOMY OF PROGRAM EVALUATION APPROACHES

The six groups used as the ordering device in this monograph were abstracted from this very rough and preliminary try to classify the types of evaluation activities that go on in adult education. This is only one way in which approaches can be classified. Others need to be developed and comparative merits debated. The point presently isn't what organizational system to use, but the need for recognizing that the term evaluation includes such a number of different phenomena that classification systems are needed to help the practitioner and the theorist develop and improve evaluative activities.

The very tentative classification system used here has several categories and subdivisions. Starting with the broadest categories first, in practice and in general discussion, there appear to be three general types of evaluation that take place in adult education:

- Evaluation of participants* (grading, certification, etc.).
- Evaluation of personnel* (merit raises, rank and promotion, Civil Service, performance reviews, etc.).
- Evaluation of programs.*

Although the three should interact closely, they often are treated as quite distinctly different fields of specialized endeavor.

Each domain has its own subcategories. Our concern currently is with the domain of program evaluation. There are several categories within it. Some are well established and well recognized; others are emerging. Program evaluation presently appears to contain the following categories:

- Evaluation of process.*
- Evaluation of results.*
- Evaluation of program as functioning, producing systems.*

Eventually program evaluation may come to mean only the third category with the other two identified as other types of evaluation.

Evaluation of process is excluded from this monograph. It would include such things as approaches to evaluating group interaction, student-teacher interaction, teacher performance, effectiveness of teaching strategy, quality of promotion, extensiveness of legitimation activities, effectiveness of use of visual aids, etc. A good deal has been done in establishing criteria and procedures for evaluating the processes and materials used in adult education.

Evaluation of results has been included because of the general pre-occupation with this type of evaluation. It has been divided into two categories:

Attainment of objectives.
Evaluation of outcomes and effects.

The first category includes the myriad of models that use objectives as the definition of what results will be judged against. The second category includes broader approaches to examining results.

Subcategorizations within evaluation of program as functioning, producing systems evolve from the content of the present models and approaches rather than from need or logic. Recent approaches to the program as a system appear to be of two kinds:

Patterns of the program system.
Patterns of the program system as mirrored in evaluation processes and data.

Within patterns of program systems, there seem to be at least two types of approaches emerging:

Evaluation of the system and its components as input into decision making.
Evaluation of effectiveness of program parts.

Patterns of program as mirrored in evaluative activities also can be further subdivided into at least two categories:

Kinds of data; types of activities.
Evaluation processes.

See Appendix Table C-1 for the composite outline of categories. As we better understand present approaches, develop new ones, and are better able to relate to program needs, it's possible that these categories will be replaced by others.

Table C-1
 A SYSTEM FOR CLASSIFYING EVALUATION APPROACHES

- I. Evaluation of participants.
- II. Evaluation of personnel.
- III. Evaluation of program.
 - A. Evaluation of process.
 - B. Evaluation of results.
 - * 1. *Results--Attainment of Objectives.*
 - * 2. *Results--Evaluation of Outcomes and Effects.*
 - C. Evaluation of programs as functioning, producing systems.
 - 1. Patterns of program systems.
 - * a. *Evaluation As Input Into Decision Making.*
 - (1) Major or general decisions.
 - (2) Specific decision situations (selection of materials, setting priorities, etc.).
 - * b. *Evaluation of Program Parts.*
 - 2. Patterns of program systems as mirrored in evaluation.
 - * a. *Evaluation--Kinds of Data; Types of Activities*
 - * b. *Evaluation Processes.*
 - (1) General processes.
 - (2) Specifically applied processes.

* Indicates the six categories that were used as major groupings for the organization of the monograph. Within level C, numerical levels were dropped to simplify the number of things the reader has to deal with.

It seems inadvisable to take space in this monograph to try to explain the various subcategories and differentiate between them. However, those familiar with some of the current models may find Appendix Table C-2 useful in exploring the categorization. It shows where I'd locate several of the emerging adult and extension education conceptualizations and some of the more well-known models from general education and health and welfare. See Table 2 on pages 44-45 for a complete listing of the approaches included in this monograph.

You will note for example, that among the adult and extension education models, Knox and Bennett both are describing ways of categorizing evaluative data. However, they're describing different things in that Knox includes a wide range of components, while Bennett concentrates only on process and outcomes. Bruce, Steele, and Kreitlow all have posed frameworks dealing with evaluation processes. However, the three are dealing with such different things that all three could be used at the same time without undue duplication.

And as you consider some of the better known conceptualizations in general education and related fields--those by Tyler, Suchman, Scriven, Stake, Stufflebeam--it's readily apparent that they're talking about quite different things. *You don't choose among them to get a particular thing done, but chooses from them according to the different things that you need to do.* Thus, you'd choose a Tyler model if you only wanted to examine the attainment of objectives; Suchman's approach if you wanted to look at results in terms of objectives plus other criteria; Scriven if you wanted to examine more results than just those specified in the objective. You'd turn to Stake or Stufflebeam if you were interested in the system involved. Stufflebeam poses the system in terms of types of evaluation which influence decision making. Stake describes system in terms of types of evaluative data that reflect the program components and the acts of evaluation.

Table C-2

LOCATION WITHIN THE CLASSIFICATION SYSTEM OF ADULT AND
EXTENSION EDUCATION AND OTHER MAJOR APPROACHES

III. Evaluation of program.

B. Evaluation of results.

1. *Results--Attainment of Objectives.*

- a. *Tylerian Models* [Tyler & others].
- b. *Instructional models* including *Goal-Referenced Instruction*. [Popham & Baker].
- c. *Program objectives approaches* including *Criteria and Objectives* [Suchman].

2. *Results--Evaluation of Outcomes and Effects.*

- a. *Goal Free Evaluation* [Scriven].
- b. *Research Models*.
- c. *Accountability*.
- d. *Multiple Change Approach* [Hayes, Jr.].

C. Evaluation of programs as functioning, producing systems.

1. Patterns of program systems.

a. *Evaluation As Input Into Decision Making.*

- (1) Major or general decisions.
 - (a) *Decision Centered Evaluation (CIPP)* [Stufflebeam; Phi Delta Kappa].
 - (b) *Differential Evaluation* [Tripodi, Fellin, & Epstein].
- (2) Specific decision situations.
 - (a) Development Models such as *IPI Formative Evaluation* [Lindvall & Cox], and *NewStart Evaluation System* [Lamrock, Smith & Warren].
 - (b) *Discrepancy Evaluation* [Provus].

b. *Evaluation of Program Plans.*

- (1) *Actual Component Approach* [Knox, Mezirow, & Darkenwald, Jr.].
- (2) *Execution-Impact Approach* [Freeman & Sherwood].
- (3) Management models such as *Program Evaluation and Review Technique (PERT)*.

2. Patterns of evaluation.

a. *Evaluation--Kinds of Data; Types of Activities.*

- (1) *Countenance of Evaluation* [Stake].
- (2) *System Role Model* [Knox].
- (3) *Means-Ends Hierarchy* [Bennett].

b. *Evaluation Processes.*

- (1) General processes.
 - (a) *Appraisal Model* [Harris].
 - (b) *Natural Process Model* [Steele].
- (2) Specifically applied processes.
 - (a) *Monitoring Evaluation* [Bruce].
 - (b) *Improvement Evaluation* [Kreitlow].

Categories can be crossed very easily with a bit of creativity. For example, although the program decision CIPP model was designed as input into decision making, each of the four types could be considered components of program (context, input, process, and product) and used as a format for judging effectiveness. On the other hand, the actual component model was designed to examine effectiveness, but with a bit of restructuring it could be cast as a guide to key decisions.

Classification systems can become extremely engrossing. They're not valuable in and of themselves. They have value only if they help better understand the phenomena being classified.

BIBLIOGRAPHY

- "Accountability in Education." *Educational Technology*, XI (January, 1971).
- Alexander, Frank D. "A Critique of Evaluation." *Journal of Cooperative Extension*, III (Winter, 1965), 205-12.
- Notes from AERA Concepts of Evaluation Workshop, Portland, Oregon, October 19-20, 1972.
- Alioto, Robert, and J. A. Jungherr. "Using PPBS to Overcome Taxpayers' Resistance." In *Emerging Patterns of Administrative Accountability*, Lesley H. Browder, Jr., ed. Berkeley, California: McCutchan Publishing Corporation, 1971, pp. 252-59.
- Alkin, Marvin C. *Towards an Evaluation Model: A Systems Approach*. Working Paper No. 4. Los Angeles: University of California, Center for the Study of Evaluation of Instructional Programs, August, 1967.
- American Institutes for Research. *Evaluative Research: Strategies and Methods*. Pittsburgh, Pennsylvania: American Institutes for Research, 1970.
- Armstrong, Robert, et al. "A Scheme for Evaluation." In *Educational Accountability Through Evaluation*, E. Wayne Roberson, ed. Englewood Cliffs, New Jersey: Educational Technology Publications, 1971.
- Axford, Roger W. *Adult Education: The Open Door*. Scranton, Pennsylvania: International Textbook Company, 1969.
- Bennett, Claude F. *Theory and Strategy for Evaluation of Knowledge Dissemination*. Washington, D.C.: Extension Service-USDA, Program and Staff Development, 1973.
- Berlak, Harold. "Values, Goals, Public Policy and Educational Evaluation." *Review of Educational Research*, XL (April, 1970).
- Blaney, John P. "The Application of Formative and Summative Evaluation Techniques to Short Term Adult Education Courses." Paper presented at the Adult Education Research Conference, Chicago, Illinois, 1972.
- Bloom, Benjamin S., ed. *Taxonomy of Educational Objectives: The Classification of Educational Goals*. Handbook 1: Cognitive Domain. New York: McKay, 1956.
- _____, J. Thomas Hastings, and George F. Madaus. *Handbook on Formative and Summative Evaluation of Student Learning*. New York: McGraw-Hill Book Company, 1971.
- Borus, Michael E., and William R. Tash. *Measuring the Impact of Manpower Programs: A Primer*. Ann Arbor, Michigan: The University of Michigan--Wayne State University, Institute of Labor and Industrial Relations, 1970.

- Boyle, Patrick G. "Criteria for Program Priorities." Paper presented at a conference of Extension Home Economists Program Leaders, Washington, D.C., November, 1972.
- _____ and Irwin R. Jahns. "Program Development and Evaluation." In *Handbook of Adult Education*, Robert M. Smith, George F. Aker, and J. R. Kidd, eds. New York: The Macmillan Company, 1970.
- Brack, Robert, "An Ecclectic Approach to Evaluation." Paper prepared for the Adult Education Research Conference, Chicago, Illinois, April, 1972.
- Browder, Leslie H., Jr. *Emerging Patterns of Administrative Accountability*. Berkeley, California: McCutchan Publishing Corporation, 1971.
- Bruce, Robert L. "Monitoring Evaluation." Ithaca, New York: Cornell University, 1972. [Mimeographed]
- Bruner, J. S. *Toward a Theory of Instruction*. Cambridge, Massachusetts: Harvard University Press, 1966.
- Byram, Harold M., and Marvin Robertson. *Locally Directed Evaluation of Local Vocational Education Programs*. East Lansing, Michigan: Michigan State University, College of Education, Department of Secondary Education and Curriculum, March, 1970.
- Byrn, Darcie, ed. *Evaluation in Extension*. Topeka, Kansas: H. M. Ives & Sons, Inc., 1959.
- Cain, Glen G., and Robinson G. Hollister. *The Methodology of Evaluating Social Action Programs*. Madison, Wisconsin: University of Wisconsin, Institute for Research on Poverty, n.d.
- Caro, Francis G., ed. *Readings in Evaluation Research*. New York: Russell Sage Foundation, 1971.
- Cass, Angelica W. *Basic Education for Adults*. New York: Association Press, 1971.
- Checklists for Public School Adult Education Programs*. Circular No. 356. Washington, D.C.: U.S. Office of Education, Federal Security Agency, n.d.
- Cohen, D. "Politics and Research: Evaluation of Social Action Programs in Education." *Review of Educational Research*, XL, 213-38.
- Committee on Evaluation. *Program Evaluation in Adult Education*. Washington, D.C.: Adult Education Association of the U.S.A., 1952.
- Corey, Stephen M. *Action Research To Improve School Practices*. New York: Columbia University, Teachers College, 1953. [Emphasizes action person doing research in addition to the professional researcher.]
- Crane, Peter, and Clark C. Abt. "A Model for Curriculum Evaluation." *Educational Technology*, IX (October, 1969), 17-25.

- Deniston, O. L., I. M. Rosenstock, W. Welch, and V. A. Getting. "Evaluation of Program Efficiency." *Public Health Reports*, LXXXIII (No. 7, 1968).
- Duft, Ken D. "Systems Planning for Extension." *Journal of Cooperative Extension*, VII (Fall, 1969), 168-78.
- Eash, J. "The Politics of Curriculum Evaluation Research." In *Curriculum Theory Network*, J. Weiss, ed. Toronto: The Ontario Institute for Studies in Education, 1972.
- Elliott, Elizabeth. "A Model for Evaluating Educational Programs Aimed at Disadvantaged Families." Paper presented at the Adult Education Research Conference, Chicago, Illinois, April, 1972.
- EMIS Data Definition Document, Fiscal Year 1974. Washington, D.C.: Extension Service-USDA, Extension Management Information System.
- Essert, Paul L. *Creative Leadership of Adult Education*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1951.
- Etzioni, Amitai. "Two Approaches to Organizational Analysis: A Critique and a Suggestion." In *Program Evaluation in the Health Fields*, Herbert C. Schulberg, Alan Sheldon, and Frank Baker, eds. New York: Behavioral Publications, Inc., 1969, pp. 101-20.
- _____ and Edward W. Lehman. "Some Dangers in 'Valid' Social Measurement." *The Annals of the American Academy*, II (September, 1967), 1-15.
- Evaluation Comment. Los Angeles: University of California, Center for the Study of Evaluation. [Periodical]
- Farmer, James A., Jr., Ralph K. Sylvester, and Patrick J. Weagraff. *Western Region AMIDS Evaluation*. Los Angeles: University of California, Division of Vocational Education, 1970.
- Forehand, Garlie A. "An Evaluation System for Curriculum Innovation." *Teacher's College Record*, LXXII (May, 1971), 576-91.
- Freeman, Howard E., and Clarence C. Sherwood. *Social Research and Social Policy*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1970.
- Furst, Edward J. *Constructing Evaluation Instruments*. New York: David McKay Co., Inc., 1964.
- Gagné, R. M. *The Conditions of Learning*. New York: Holt, Rinehart and Winston, 1956.
- Glass, Gene V. "Educational Product Evaluation: A Prototype Format Applied." *Educational Researcher*, I (January, 1972), 7-10.
- _____. *The Growth of Evaluation Methodology*. AERA Monograph Series on Curriculum Evaluation. Chicago: Rand McNally, 1972.

- _____ and B. Worthen. "Educational Evaluation and Research: Similarities and Differences." In *Curriculum Theory Network*, J. Weiss, ed. Toronto: The Ontario Institute for Studies in Education, 1972.
- Gottman, John Mordechai, and Robert Earl Clasen. *Evaluation in Education: A Practitioner's Guide*. Itasca, Illinois: F. E. Peacock Publishers, Inc., 1972.
- Guba, E. G. "The Failure of Educational Evaluation." *Educational Technology*, IX (1969), 29-38.
- Hale, Larry A. "Evaluation Is a Process of Deciding the Value of Something." Paper prepared for the Missouri Extension Program Development Conference, 1970.
- Harris, Chester W. "The Appraisal of a School--Problem for Study." *Journal of Educational Research*, XLI (November, 1947), 172-82.
- Hammond, Robert L. *Evaluation at the Local Level*. Columbus, Ohio: Ohio State University Evaluation Center, 1967.
- Hawkrige, D. G. "Designs for Evaluative Studies in American Institute for Research." In *Evaluative Research--Strategies and Methods*. Pittsburgh, Pennsylvania: American Institutes for Research, 1970.
- Hayes, Samuel P., Jr. *Evaluating Development Projects*. Place de Fontenoy, Paris: United National Education, Scientific & Cultural Organization, 1959.
- Hemphill, J. K. "The Relationship Between Research and Evaluation Studies." In *Educational Evaluation: New Roles, New Means*, R. W. Tyler, ed. Chicago: National Society for the Study of Education, 1969.
- Herzog, Elizabeth. *Some Guidelines for Evaluative Research*. Washington, D.C.: U.S. Department of Health, Education, and Welfare, Social Security Administration, 1959.
- Hesseling, P. *Strategy of Evaluation Research*. The Netherlands: Royal Van Corcum Ltd., pp. 217-34, as cited in *A Strategy for Evaluation Design*, Casper F. Paulson, Jr., ed. Oregon State System of Higher Education, 1970, section VII, pp. 22-24.
- Honnold, Robert E., et al. "Comprehensive Reviews." Materials prepared by a Program and Staff Development Task Force, Extension Service-USDA, 1972.
- Houle, Cyril O. *The Design of Education*. San Francisco: Jossey Bass, Inc., 1972.
- "Impact Evaluation." Materials prepared as a follow-up of a University of Wisconsin-Extension Workshop on Impact Evaluation, Madison, Wisconsin, March, 1972.
- Kaplan, Abraham. *The Conduct of Inquiry*. San Francisco: Chandler Publishing Company, 1964.

- Kempfer, Homer. *Adult Education*. New York: McGraw-Hill Book Company, Inc., 1955, pp. 404-19.
- Kirkpatrick, Donald L. "Evaluation of Training." In *Training and Development Handbook*, Robert L. Craig and Lester R. Bittel, eds. New York: McGraw-Hill Book Company, 1967, pp. 87-112.
- Knowles, Malcolm S. *The Modern Practice of Adult Education*. New York: Association Press, 1970.
- Knox, Alan B. "Continuous Program Evaluation." In *Administration of Continuing Education*, Nathan C. Shaw, ed. Washington, D.C.: National Association for Public School Education, 1969.
- _____. "Improving Urban School Adult Education: An Analysis of the Adult and Community Education Programs of the Washington, D.C. Schools." New York: Columbia University, Teachers College, 1967.
- _____. *Program Evaluation in Adult Basic Education*. Tallahassee, Florida: Florida State University, Adult Education Research Information, Processing Center, 1971.
- _____, Jack Mezirow, and Gordon G. Darkenwald. *Evaluating Adult Basic Education: A Strategy and a Manual*. New York: Columbia University, Teachers College, Center for Adult Education, 1973 [pre-publication draft].
- Krathwohl, D. R., B. S. Bloom, and B. B. Masia. *Taxonomy of Educational Objectives: The Classification of Educational Goals*. Handbook 2: Affective Domain. New York: McKay, 1964.
- Kreitlow, Burton W. *Indiana Adult Basic Education--An Improvement Evaluation*. Indianapolis, Indiana: State Department of Public Instruction, 1972.
- Kropp, Russell P., and Coolie Verner. "An Attitude-Scale Technique for Evaluating Meetings." *Adult Education*, VII (Summer, 1957), 212-15.
- Lamrock, A. L., A. D. Smith, and P. W. Warren. "Evaluation: Its Scope and Systems for Evaluation Development." Paper prepared for Meeting of Research Directors of NewStart Corporation, Ottawa, Canada, March 29-31, 1971.
- Lange, Robert R. "Evaluation As Management Methodology." In *A Strategy for Evaluation Design*, Casper F. Paulson, Jr., ed. Oregon State System of Higher Education, 1970, Section VII.
- Lewis, Robert B. "Motivation Model for Extension." *Journal of Extension*, X (Winter, 1972), 23-24.
- Lindvall, C. M., and Richard Cox. *The IPI Evaluation Program*. AERA Monograph Series on Curriculum Evaluation. Chicago: Rand McNally & Company, 1970.
- Longest, James W. "Designing Evaluative Research." Paper presented at the annual meeting of the Rural Sociological Society, San Francisco, California, 1969.

- Loomis, C. P. *Social Systems: Essays on Their Persistence and Changes*. Princeton, New Jersey: Van Nostrand, 1960.
- McGuire, Thomas O. "Decisions and Curriculum Objectives: A Methodology for Evaluation." *Alberta Journal of Educational Research*, XV (March, 1969), 17-30.
- Mager, Robert F. *Goal Analysis*. Felmont, California: Fearon Publishers/Lear Siegler, Inc., Educational Division, 1972.
- _____. *Preparing Instructional Objectives*. Felmont, California: Fearon Publishers/Lear Siegler, Inc., Educational Division, 1962.
- Matteson, Harold R. *The Process of Evaluating Vocational Educational Programs in Agriculture*. Instructional Series No. 3. Madison, Wisconsin: University of Wisconsin, Department of Agricultural and Extension Education, February, 1972.
- Miller, Harry L., and Christine H. McGuire. *Evaluating Liberal Adult Education*. Chicago: Center for the Study of Liberal Education for Adults, 1961.
- Merwin, Jack C., and Frank B. Womer. "Evaluation in Assessing the Progress of Education To Provide Bases of Public Understanding and Public Policy." In *Educational Evaluation: New Roles, New Means*, Ralph Tyler, ed. Chicago: The University of Chicago, 1970, pp. 306-12.
- Munson, Byron E. *Changing Community Dimensions*. Columbus, Ohio: Ohio State University, 1968.
- Nelson, Orville. "The Evaluation System for the American Industry Secondary School Courses." Paper presented at the AVERA program at the AVA Convention in Dallas, Texas, December, 1968.
- Novick, David, ed. *Program Budgeting*. Cambridge, Massachusetts: Harvard University Press, 1965.
- Paulson, Belden. *A Model for Community Analysis: Steps in Planning the Total Health of a Community*. Milwaukee, Wisconsin: University of Wisconsin-Extension, Center for Leadership Development, April, 1972.
- Peterson, Richard E. *The Crisis of Purpose: Definition and Uses of Institutional Goals*. Princeton, New Jersey: Educational Testing Service, 1971.
- Phi Delta Kappa National Study Committee on Evaluation. *Education and Decision Making*. Itasca, Illinois: F. E. Peacock Publishers, Inc., 1971.
- Popham, W. James. *An Evaluation Guidebook*. Los Angeles: The Instructional Objectives Exchange, 1972.
- _____, and Eva L. Baker. *Systematic Instructions*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1971.
- _____, Elliot W. Eisner, Howard J. Sullivan, and Louise L. Tyler. *Instructional Objectives*. AERA Monograph Series on Curriculum Evaluation, No. 3. Chicago: Rand McNally & Company, 1969.

Provus, Malcolm. *Discrepancy Evaluation*. Berkeley, California: McCutchan Publishing Co., 1971.

_____. "Evaluation as Public Policy." In *Curriculum Theory Network*, J. Weiss, ed. Toronto: The Ontario Institute for Studies in Education, 1972.

Raudabaugh, J. Neil. "Evaluation As a Concept." Presentation to the ECOP-EMIS Advisory Committee Meeting, Lake Texoma, Oklahoma, November, 1970.

Rippey, Robert M. "Can Evaluation Be Integral to Innovation?" In *Curriculum Evaluation: Potentiality and Reality*, Joel Weiss, ed. Curriculum Theory Network Monograph Supplement. Ontario, Canada: Ontario Institute for Studies in Education, 1972, pp. 45-58.

_____. "Introduction: What Is Transactional Evaluation?" Paper presented at the meeting of the American Educational Research Association in Chicago, 1972.

Sabrosky, Laurel. "Evaluation." In *The Cooperative Extension Service*, H. C. Saunders et al. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1966, pp. 339-51.

Sanders, H. C. *Instruction in the Cooperative Extension Service*. Baton Rouge, Louisiana: Louisiana State University and Agricultural and Mechanical College, Publishers, 1972, pp. 76-87.

Scriven, Michael S. "The Methodology of Evaluation." In *Perspectives of Curriculum Evaluation*, Ralph W. Tyler, Robert Gagné, and Michael Scriven, eds. AERA Monograph Series on Curriculum Evaluation, No. 1. Chicago: Rand McNally & Co., 1967.

_____. "Prose and Cons About Goal-Free Evaluation." *Evaluation Comment* (December, 1972), III.

Seaman, Don F., and Donnie Dutton. "Self Evaluation." In *Materials and Methods in Adult Education*, Chester Klevins, ed. Canoga Park, California: Klevins Publications, Inc., 1972, pp. 265-74.

Schulberg, Herbert C., and Frank Baker. "Program Evaluation Models and the Implementation of Research Findings." In *Program Evaluation in the Health Fields*, Herbert C. Schulberg, Alan Sheldon, and Frank Baker, eds. New York: Behavioral Publications, Inc., 1969, pp. 562-72.

Stake, Robert E. "The Countenance of Educational Evaluation." *Teachers College Record*, LXVIII (1967), 523-40.

_____. "Objectives, Priorities and Other Judgment Data." *Review of Educational Research*, XL (April, 1970), 181-212.

_____. "A Research Rationale for EPIE." *The EPIE Forum*, I (September, 1967).

_____. and Terry Denny. "Needed Concepts and Techniques for Utilizing More Fully the Potential of Evaluation." In *Educational Evaluation: New Roles and New Means*, Part II, 68th Yearbook of the National Society for the Study of Education, Ralph W. Tyler, ed. Chicago: The University of Chicago Press, 1969, chapter XVI, pp. 370-90.

Starr, Harold, et al. "A System for State Evaluation of Vocational Education." Research Series No. 58. Washington, D.C.: U.S. Department of Health, Education and Welfare, Office of Education, Bureau of Research, May, 1970.

Steele, Sara M. "A Contemporary Concept of Program Evaluation." Madison, Wisconsin: University of Wisconsin-Extension, 1972. [Mimeographed]

_____. *Cost-Benefit Analysis and the Adult Educator*. Washington, D.C.: ERIC Clearinghouse on Adult Education and Adult Education Association of the U.S.A., 1971.

_____. "Determining the Quality, Importance and Suitability of Programs." Madison, Wisconsin: University of Wisconsin, Department of Agricultural and Extension Education, 1970.

_____. "Exploring Participants Views of Programs." Material prepared for students in a course on Evaluation in Extension, University of Wisconsin-Madison, Spring, 1972.

_____. "Program Evaluation--A Broader Definition." *Journal of Extension*, VIII (Summer, 1970), 5-17.

_____. "Program Evaluation As an Administrative Concept." Paper presented at AERA annual meeting, New Orleans, February 28, 1973.

_____. *Programmer's Responsibility in Examining Program Effectiveness*. Evaluation Series. Madison, Wisconsin: University of Wisconsin-Extension, Division of Program and Staff Development, October, 1972. [Identifies program personnel responsibility regardless of whether help is gotten from evaluation or research specialists.]

_____. *Six Dimensions of Program Effectiveness*. Madison, Wisconsin: University of Wisconsin-Extension, Division of Program and Staff Development, 1972.

_____ and Robert E. Brack. "Evaluating the Attainment of Objectives: Process, Problems and Prospects." [Manuscript in process.]

Stufflebeam, Daniel L. "Evaluation As Enlightenment for Decision-Making." Address at working conference on assessment theory, the Commission of Assessment of Educational Outcomes, the Association for Supervision and Curriculum Development, Sarasota, Florida, January, 1968.

_____. "The Relevance of the CIPP Evaluation Model for Educational Accountability." *Journal of Research and Development in Education*, V (February, 1971), 19-25.

Suchman, Edward A. *Evaluative Research*. New York: Russell Sage Foundation, 1967.

Taba, Hilda, and Enoch Sawin. "A Proposed Model of Evaluation." *Educational Leadership*, XX (October, 1962), 57-59ff.

- Taylor, Peter A., and Thomas O. McGuire. "A Theoretical Evaluation Model." *Manitoba Journal of Educational Research*, I (1966), 12-17.
- Thiede, Wilson. "Evaluation and Adult Education." In *Adult Education: Outlines of an Emerging Field of University Study*, Gale Jensen, A. A. Liveright, and Wilbur Hallenbeck, eds. Washington, D.C.: Adult Education Association, 1964.
- Toward a Social Report*. Washington, D.C.: U.S. Department of Health, Education and Welfare, January, 1969.
- Tripodi, Tony, Phillip Fellin, and Irwin Epstein. *Social Program Evaluation*. Itasca, Illinois: F. E. Peacock Publishers, Inc., 1971.
- Tyler, Ralph W. *Basic Principles of Curriculum and Instruction*. Chicago: University of Chicago Press, 1950.
- Uhl, Norman P. *Identifying Institutional Goals*. Durham, North Carolina: National Laboratory for Higher Education, 1971.
- Van Gigch, John P., and Richard E. Hill. *Using Systems Analysis to Implement Cost Effectiveness and Program Budgeting in Education*. Englewood Cliffs, New Jersey: Educational Technology Publications, 1971.
- Warner, W. Keith. "Problems in Measuring the Goal Attainment of Voluntary Organizations." *Journal of Adult Education*, XIX (No. 1, 1967).
- Wedemeyer, Charles A. "Evaluation of Continuing Education Programs" *American Journal of Pharmaceutical Education*, XXXIII (No. 5), 793-863.
- Weiss, Carol H. *Evaluation Research*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1972.
- Wholey, Joseph, et al. *Federal Evaluation Policy: Social Evaluation by Federal Agencies*. Washington, D.C.: The Urban Institute, 1970.
- Wilhelms, Fred T., ed. *Evaluation As Feedback and Guide*. Washington, D.C.: Association for Supervision and Curriculum Development, NEA, 1967.
- Wilson, Meredith C., and Gladys Gallup. *Extension Teaching Methods*. Extension Service Circular 495. Washington, D.C.: Federal Extension Service-USDA, August, 1955.
- Wittrock, M., and D. Wiley. *Evaluation of Instruction: Issues and Problems*. New York: Holt, Rinehart and Winston, 1970.
- Young, Stanley. "Organization As a Total System." In *Emerging Patterns of Administrative Accountability*, Leslie H. Browder, Jr., ed. Berkeley, California: McCutchan Publishing Corporation, 1971, pp. 311-30.

INDEX OF AUTHORS

- Abt, Clark C., 96-97, 233
- Alexander, Frank D., 131-33, 233
- Alioto, Robert, 178-79
- Alkin, Marvin C., 110-11, 227
- Armstrong, Robert, 180, 221
- Axford, Roger W., 161
- Baker, Eva L., 164-68, 225
- Baker, Frank, 117, 229
- Bennett, Claude F., 129-30, 227
- Berlak, Harold, 205-206, 231
- Bloom, Benjamin S., 168-70, 184-85, 224
- Borus, Michael E., 203-204, 225
- Boyle, Patrick G., 86-88, 160, 230
- Brack, Robert E., 54, 186-88
- Browder, Leslie H., Jr., 217-19
- Bruce, Robert L., 146-47, 227
- Bruner, J. S., 184-85
- Byram, Harold M., 176-77, 226
- Byrn, Darcie, 159
- Cain, Glen G., 212-14, 231
- Cass, Angelica W., 161
- Clasen, Robert, 156, 225
- Cook, Desmond L., 107-108
- Cox, Richard, 89-91, 226
- Crane, Peter, 96-97, 233
- Darkenwald, Gordon G., Jr.,
104-106, 220
- Deniston, O. L., 173-74, 229
- Denny, Terry, 120-23
- Duft, Ken D., 112-14, 230
- Dutton, Donnie, 161
- Elliott, Elizabeth, 161
- Eqstein, Irwin, 81-83, 223
- Essert, Paul L., 162
- Etizioni, Amitai, 117, 229
- Farmer, James A., Jr., 161
- Fellin, Phillip, 81-83, 223
- Forehand, Garlie A., 207-209, 226
- Freeman, Howard E., 102-103, 224
- Furst, Edward J., 184-85
- Gagne, R. M., 184-85
- Gallup, Gladys, 202, 224
- Glass, Gene V., 95-96, 232
- Gottman, John M., 156, 225
- Hale, Larry A., 161
- Hammond, Robert L., 180, 221
- Harris, Chester W., 135-37, 220
- Hastings, J. Thomas, 168-70, 224
- Hayes, Samuel P., Jr., 198-99, 228
- Hesseling, P., 115, 224
- Hill, Richard E., 111-12, 232
- Hollister, Robinson G., 212-14, 231
- Houle, Cyril O., 161, 188-89
- Jahns, Irwin R., 160
- Jungherr, J. A., 178-79

- Kempfer, Homer, 162
- Knowles, Malcolm, 161
- Knox, Alan B., 104-106,
124-28, 220, 232
- Krathwohl, D. R., 184-85
- Kreitlow, Burton W., 148-49, 225
- Lamrock, A. L., 91-94, 195-97, 229, 233
- Lewis, Robert B., 118, 227
- Lindvall, C. M., 89-91, 226
- Longest, James W., 212-14, 231
- Loomis, C. P., 116-17, 231
- Madaus, George F., 168-70, 224
- Mager, Robert F., 184-85
- Masia, B. B., 184-85
- Matteson, Harold R., 174-76, 221
- McGuire, Christine H., 158
- McGuire, Thomas O., 186
- Merwin, Jack C., 155-56
- Mezirow, Jack D., 104-106, 220
- Miller, Harry L., 158
- Nelson, Orville, 181, 226
- Novick, David, 178-79
- Paulson, Belden, 210-11, 231
- Phi Delta Kappa, 78-80, 138-42, 223
- Popham, W. James, 164-68, 225
- Provus, Malcolm, 83-84, 223
- Riphey, Robert M., 150-51, 232
- Raudabaugh, J. Neil, 161
- Robertson, Marvin, 176-77, 226
- Roman, Daniel D., 107-108
- Ryden, Einar R., 107-08
- Sabrosky, Laurel, 159-60
- Sanders, H. C., 160-61
- Sawin, Enoch, 179-80, 229
- Schulberg, Herbert C., 117, 229
- Scriven Michael S., 192-94, 225
- Seaman, Don F., 161
- Sherwood, Clarence C., 102-03, 224
- Smith, A. D., 91-94, 195-97, 229, 233
- Stake, Robert E., 120-23, 221
- Starr, Harold, 177-78, 229
- Steele, Sara M., 98-100, 143-45, 186-88,
200-201, 215-16, 228
- Stufflebeam, Daniel L., 78-80, 223
- Suchman, Edward A., 171-73, 186, 221
- Sylvester, Ralph K., 161
- Taba, Hilda, 179-80, 229
- Tash, William R., 203-204, 225
- Taylor, Peter A., 186
- Thiede, Wilson, 157-58
- Tripodi, Tony, 81-83, 223
- Tyler, Ralph W., 154-55
- Van Gigch, John P., 111-12, 232
- Warner, W. Keith, 188-89

Warren, P. W., 91-94, 195-97, 229, 233

Weagraff, Patrick J., 161

Wedemeyer, Charles A., 161

Weiss, Carol H., 212-14, 231

Wholey, Joseph, 181-83, 233

Wilson, Meredith C., 202, 224

Womer, Frank B., 155-56

Young, Stanley, 109-10, 230

INDEX OF APPROACHES

<u>Title</u>	<u>Classification Number</u>	<u>Author</u>	<u>On page:</u>
Accountability	6.12		217
Actual Component Approach	2.2	Knox, Mezirow, & Darkenwald, Jr.	104
Adoption of the Tyler Model by Adult Education	5.1		157
Appraisal Model	4.1	Harris	135
Characteristics of Objectives	5.4		188
Countenance of Evaluation	3.1	Stake	120
Criteria and Objectives	5.3	Matteson	174
Criteria of Success Approach	5.3	Suchman	171
Cube Models	5.3	Hammond; Armstrong <u>et al.</u>	180
Data Management	4.2	Phi Delta Kappa	138
Decision Centered Evaluation (CIPP)	1.1	Stufflebeam; Phi Delta Kappa	78
Developmental Evaluation	1.5		89
Differential Evaluation	1.2	Tripodi, Fellin, & Epstein	81
Discrepancy Evaluation	1.3	Provus	84
Effectiveness of Methods	6.5	Wilson & Gallup	202
Efficiency Examinations	6.11		215
Evaluating Decision Making	2.3	Hesseling	115
Evaluation As Facilitation of Learning	5.2	Bloom, Hastings, & Madaus	168
Execution-Impact Approach	2.1	Freeman & Sherwood	102
Four-Question Approach	5.1	Gottman & Clasen	156

<u>Title</u>	<u>Classification Number</u>	<u>Author</u>	<u>On page:</u>
Goal Free Evaluation	6.1	Scriven	192
Goal-Referenced Instruction	5.2	Popham & Baker	164
Impact Evaluation	6.6	Borus & Tash	203
Improvement Evaluation	4.5	Kreitlow	148
Information Domains	5.3	Nelson	181
Institutional Evaluation	6.8	Forehand	207
Instructional Evaluation Approaches	5.2		164
IPI Formative Evaluation	1.5	Lindvall & Cox	89
Kinds of Objectives Operating Within a Program	5.4		186
Level and Time Sequence of Objectives	5.4		186
Locally Directed Evaluation	5.3	Byram & Robertson	176
Macro System Model	2.3	Alkin	110
Management Information Systems (SEMIS)	2.3		114
Managerial Systems	2.3		107
Materials Evaluation	1.6		95
Means-Ends Hierarchy	3.3	Bennett	129
Monitoring Evaluation	4.4	Bruce	146
Motivational Model	2.4	Lewis	118
Multiple Change Approach	6.3	Hayes, Jr.	198
Multiple Dimensions of Program Effectiveness	6.4	Steele	200
National Assessment Program	5.1		155
Natural Process Approach	4.3	Steele	143
NEA Model	5.3	Taba & Sawin	179
NewStart Evaluation System	1.5	Lamrock, Smith, & Warren	91

<u>Title</u>	<u>Classification Number</u>	<u>Author</u>	<u>On page:</u>
O-A-R Model	5.3	Deniston <u>et al.</u>	173
Objectives As Systems and Parts of Systems	5.4		184
Objectives: Categorizations and Criteria	5.4		184
Ohio Model	5.3	Starr <u>et al.</u>	177
Organizational Models	2.4	Etizioni; Schulberg & Baker	117
Organization As a Total System	2.3	Young	109
The Original Tyler Model	5.1	Tyler	154
Participant Reaction Approaches	1.7		98
Priority Decisions	1.4	Boyle	86
Program Contact System	2.4	Duft	112
Program Evaluation and Review Technique (PERT)	2.3		107
Program Objectives Approaches	5.3		171
Program Planning Budgeting System (PPBS)	5.3		178
Public Policy	6.7	Berlak	205
Research Models	6.10	Weiss; Longest; Cain & Hollister	212
Social Indicators	6.9	Paulson	210
Social Systems Models	2.4	Loomis	116
Socio-Organizational Systems	2.4		116
System Approach to Goal Setting	2.3	Van Gigch & Hill	111
System Role Model	3.2	Knox	124

<u>Title</u>	<u>Classification Number</u>	<u>Author</u>	<u>On page:</u>
System Variables Approaches	5.3		179
Trade-Off and Comparative Cost Approach	1.6	Glass	95
Transactional Evaluation	4.6	Rippey	150
Tylerian Models	5.1		154
Types of Evaluation	5.3	Wholey <u>et al.</u>	181
Weighted Criteria Approach	1.6	Crane & Abt	96
When-To-Do-It-Yourself Continuum	3.4	Alexander	131
Zones of Results	6.2	Lamrock, Smith, & Warren	195

